Giovanni Rolla

FILOSOFIA

uma introdução temática

nel

FILOSOFIA

uma introdução temática

Giovanni Rolla

FILOSOFIA

uma introdução temática

Florianópolis
NEL
2021

para Ana e Jorge, por tudo

Sumário

Prefácio

A escrita deste livro me manteve intelectualmente engajado enquanto confinado durante a pandemia do novo coronavírus. Com efeito, o isolamento permitiu-me realizar um projeto que eu havia concebido alguns meses antes ao constatar que minhas aulas de Introdução à Filosofia na Universidade Federal da Bahia reuniam um material com uma abordagem original e instigante, ou pelo menos eu assim espero. Minha experiência anterior com um livro introdutório — daquela vez, especificamente sobre epistemologia — foi amplamente positiva, o que me motivou a escrever novamente para um público não especializado. Desta vez, no entanto, a tarefa mostrou-se bem mais complicada por tratar-se de uma perspectiva filosófica mais geral. Em virtude das circunstâncias, na medida em que escrevi, adaptei, excluí e ampliei muitas discussões, de modo que este livro não é um retrato fiel das minhas aulas, mas apenas vagamente inspirado nelas.

Com o objetivo de facilitar a leitura para quem não é acostumado com a redação acadêmica, eu tentei manter o mesmo modelo de escrita com que havia tido algum sucesso no outro livro, isto é, sem referências no corpo do texto e sem notas de rodapé. Logo notei que esse modelo, apesar de garantir certa fluidez, limitava a apresentação das discussões filosóficas, porque me impedia de expressar adequadamente a dialética dessas discussões. Eu cedi, mas não inteiramente: fui aos poucos incluindo menções aos filósofos e às filósofas cujos trabalhos são indispensáveis para os debates apresentados aqui. Ainda assim, no entanto, preservo a mesma postura essencialmente

temática do livro anterior, pois sempre me pareceu menos importante saber quem disse o que do que ser capaz de avaliar se aquilo que é dito é verdadeiro ou ao menos bem arrazoado. Com isso eu não nego que a exegese e a compreensão histórica sejam instrumentos importantes e valiosos para um trabalho acadêmico sério, mas eu enfaticamente nego que a mera compreensão de um autor e da sua situação história deva ser a *finalidade* de um inquérito filosófico. Por isso, este livro busca primariamente ser um exemplo de investigação (em um nível introdutório e meramente elementar) sobre problemas filosóficos, não sobre o que pensaram filósofos e filósofas. É inclusive possível que eu tenha cometido alguns graves deslizes interpretativos. Tudo bem: se você suspeita que algum filósofo ou filósofa não disse exatamente aquilo que eu lhe atribuí, coloque mentalmente um asterisco ao lado de seu nome a cada nova ocorrência. Assim, Filósofo(a)* disse o que Filósofo(a) não disse, mas talvez devesse ter dito.

Quando questionei meu amigo Nykolas Friederich von Petters sobre se seria um prejuízo datar este trabalho com tantas referências às crises pelas quais passamos, ele me confortou com a declaração enigmática, como é típico de filósofos com nomes tão europeus, de que não seria um problema porque "a praga é uma condição da humanidade". Isso me persuadiu, embora eu ainda não tenha entendido exatamente o que ele havia querido dizer. Além disso, ao ponderar exaustivamente sobre se deveria insistir em referências que poderiam alienar leitores vindos de tempos mais bem afortunados, conclui que, embora este livro possa tornar-se datado pelo plano de fundo, talvez seja interessante um dia poder olhar para este trabalho como um produto de seu tempo, como uma peça histórica, o que admitidamente não é muito comum na minha tradição.

Eu digo isso porque fui treinado como um filósofo analítico, o que transparece na seleção e na abordagem dos temas. Portanto, se você tem em mãos este livro com a expectativa de encontrar discussões esclarecedoras sobre fenomenologia, hermenêutica ou estética, infelizmente, você se frustrará (e a culpa é toda minha). Esses não são temas tão tradicionais na filosofia analítica quanto filosofia da linguagem e

filosofia da ciência, por exemplo — e, além disso, eu não teria nada inteligente a dizer, nada sequer que merecesse ser dito sobre aqueles assuntos que são tipicamente associados à filosofia continental. Eu confio que existam na nossa língua introduções muito qualificadas às quais você deve recorrer se quiser saber mais sobre aqueles temas. Enfim, admito que talvez o livro deveria chamar-se 'Filosofia *analítica*: uma introdução temática', mas acho que não teria tanto apelo — e, acima de tudo, eu confesso que não quero pregar para meu rebanho, mas converter os desavisados.

Eu tenho muitas pessoas a agradecer por este livro. Conforme o escrevia, percebia que muitas das minhas escolhas remetiam às aulas brilhantes e extremamente cativantes de Introdução à Filosofia que Paulo Francisco Estrella Faria ministrava na Universidade Federal do Rio Grande do Sul em meados dos anos 2000. Eu tive a sorte de ter sido seu aluno logo desde o começo da minha formação acadêmica, principal razão pela qual continuei estudando filosofia durante minha graduação. Agradeço a Paulo pela inspiração e, além disso, pela rigorosa formação analítica, que só não foi mais rigorosa por desleixo meu. Algumas discussões que aparecem neste livro, sobretudo em filosofia da linguagem, são flagrantemente reminiscentes dos instigantes Seminários Avançados de Pesquisa também ministrados por Paulo nas saudosas tardes de sexta-feira, de 2008 a 2010. Sou, pois, triplamente grato a Paulo.

Também agradeço aos meus amigos e colegas Gregory Gaboardi, Marta Mangiarulo, Adriano Nunes, Gilberto Morbach, Lucas Bispo e Rafael Azize pelos seus valiosos comentários, todos cruciais para o refinamento deste material e essenciais para que eu continuasse motivado a escrevê-lo durante um período tão difícil para nós todos. Sem esse retorno tão qualificado, este seria apenas a sombra de um livro.

Por fim, eu sou imensamente, eternamente grato à minha querida Moiara, não apenas por ter lido cuidadosamente cada linha deste livro antes de todas as outras pessoas e por tê-lo discutido extensivamente comigo, compartilhando assim um pouco da sua visão que elevou clareza, a relevância e a profundidade deste trabalho, mas tam-

bém por ter sido uma companhia inspiradora durante a quarentena, por ter me encorajado quando eu achava que não seria competente o suficiente para finalizar este projeto, por ter me incentivado a mudar o que deveria ser mudado e a manter o que deveria ser mantido, e por ser uma interlocutora atenta e brilhantemente crítica. Eu agradeço também a Klaus e a Ostara, nossos dois gatinhos, que nos acordaram cedo todas as manhãs para que eu logo começasse a trabalhar.

Introdução

Se você tem este livro em mãos, ou — o que seria mais ambientalmente consciente — este pdf em tela, provavelmente você tem alguma curiosidade, ainda que difusa e nebulosa, sobre a filosofia. Mas também pode ser um mero acaso, e talvez você jamais tenha pensado sobre filosofia, o que quer que isso seja. De qualquer modo, esta é minha chance de cativar-lhe, pelo menos pelas próximas linhas. Então preste atenção, e eu prometo que vou tentar ser breve.

A pergunta que eu imagino que você esteja se fazendo é: *o que é a filosofia?* Admitidamente, essa é uma pergunta incômoda para quase todos os filósofos. Nós — eu e os demais filósofos (digo isso sem nenhuma vergonha, até porque não há nada de especial com o título de filósofo, embora ele tenha sido usurpado por algumas pessoas que, ao se apresentarem como tais, pretendem transmitir uma imagem de misticismo e de sabedoria oculta) — eu retomo: nós, os filósofos, não temos uma resposta consensual e incontroversa a essa pergunta. Não parece especialmente curioso que, apesar de a filosofia ser uma das atividades intelectuais mais antigas da humanidade, os filósofos discordem entre si sobre o que exatamente é filosofia? É de se esperar, por exemplo, que engenheiros concordem sobre o que é a engenharia, ou que médicos concordem sobre o que é a medicina. Claro, talvez seja o caso que eles tenham outras coisas mais urgentes para fazer do que pensar sobre a natureza de suas disciplinas. Mas nós levamos a sério esse tipo de questão, e a questão sobre o que é a filosofia não poderia ser exceção.

Neste livro, nós veremos duas respostas à pergunta sobre a natureza da filosofia: a resposta que eu chamo aqui de tradicional, porque ela é herdada da modernidade e foi aceita pela maioria dos filósofos desde então. A visão tradicional de filosofia entende que essa seja uma atividade de exame e de esclarecimento dos conceitos através dos quais pensamos a realidade. Ou seja, a filosofia consistiria em *pensar sobre o pensamento*, por assim dizer. A outra resposta possui representantes por toda história da filosofia, mas ganha maior expressão e apelo na segunda metade do século XX. Segundo essa alternativa, a filosofia é uma atividade investigativa contínua com as ciências empíricas, isto é, ciências que dependem de experimentação (como a biologia, a química, algumas áreas da física, etc.). Ou seja, a filosofia seria, assim como essas ciências, uma *investigação sobre a realidade ela mesma*.

Enquanto examinamos essas respostas e as razões pelas quais alguns pensadores conceberam a filosofia de uma maneira ou de outra, nós vamos fazer filosofia. Ou seja, estaremos examinando, contestando e atestando a razoabilidade de distinções, argumentos, teses, problemas e soluções filosóficas. Preciso adverti-los de que nosso interesse ao fazermos isso não é histórico nem exegético: não nos interessa primariamente saber quem disse o que, quem veio antes de quem, quem respondeu para quem. Há indubitavelmente obras muito qualificadas para entendermos a história da filosofia, e este livro claramente não se propõe a ser uma delas. Nosso foco aqui é *temático*: queremos entender, de uma perspectiva introdutória, elementos de *temas* filosóficos, como a filosofia da linguagem, a filosofia da ciência e a ética. Um conhecimento mais aprofundado sobre cada um desses temas requer a leitura de materiais dedicados a eles, e eu faço algumas recomendações nesse sentido ao final de cada capítulo.

No primeiro capítulo, é apresentada uma concepção consensual de filosofia como uma atividade crítica que visa examinar e resolver certas perplexidades fundamentais, perplexidades essas que são contínuas com os nossos inquéritos de dia-a-dia. A filosofia é vista como semelhante em alguns aspectos e diferente em outros do que chama-

mos, sem preconceito, de *senso comum*. Da mesma maneira, há algo em comum entre ciência e filosofia, assim como parece haver algo de substancialmente diferente entre esses dois tipos de atividade.

No segundo capítulo, exploramos a suposta diferença entre ciência e filosofia através de três distinções clássicas, sejam elas: entre conhecimentos *a priori* e *a posteriori*, entre enunciados analíticos e sintéticos e entre verdades necessárias e contingentes (não se preocupe com entender todos esses termos agora, esse entendimento virá mais adiante conforme avançamos). Essas distinções servem para justificar a ideia de filosofia como uma atividade de pensar sobre o pensamento — ou, como eu chamo, uma disciplina de segunda-ordem (por razões que ficarão claras mais adiante).

No terceiro capítulo, vemos críticas às distinções analítico/sintético e *a priori*/*a posteriori*. Também vemos argumentos persuasivos para as possibilidades de conhecer *a priori* verdades contingentes e de conhecer *a posteriori* verdades necessárias. Essas críticas abalam os fundamentos da ideia de que a filosofia é um pensar sobre o pensamento e fazem-nos reconsiderar as relações aparentemente óbvias entre as distinções apresentadas no capítulo anterior. Especialmente as críticas às distinções entre analítico e sintético e *a priori* e *a posterior* servem de motivação para ideia de que a filosofia não é substancialmente diferente das ciências empíricas.

No capítulo quatro, tratamos de noções como *verdade*, *argumentação*, *falácias* e *vieses* (entre outras noções), o que eu considero ser o mínimo necessário de lógica para, por assim dizer, "molhar os pés" nas águas profundas dos problemas filosóficos. O itinerário pode parecer estranho, porque geralmente começa-se com noções de lógica para depois avançar às questões filosóficas mais substanciais (como se a lógica não fosse substancial o bastante). A minha escolha pelo caminho contrário está baseada na hipótese de que um leitor ou uma leitora que tenha em mãos este livro já possui alguma familiaridade com argumentação e até mesmo com uma noção intuitiva e incontroversa de verdade, o que permite traçar o percurso previsto até chegar nas

discussões sobre filosofia da linguagem e sobre lógica propriamente dita (ainda que de um ponto majoritariamente informal).

No quinto capítulo, tratamos de filosofia da ciência e de um fenômeno que eu chamo de *crise de racionalidade*. Nossos problemas neste capítulo envolvem, por exemplo, a dificuldade de apresentar um critério que permita diferenciar boa ciência de asneiras como o terraplanismo e o movimento antivacina. Também examinamos algumas tentativas de caracterização do método típico da ciência, como o verificacionismo, o indutivismo e o falsificacionismo. Fechamos esse capítulo com a questão sobre se a ciência gera mesmo conhecimento.

No sexto e último capítulo, discutimos acerca de noções centrais da ética, como *certo* e *dever*. Examinamos primeiro as teorias que identificam aquilo que é moralmente correto com o que é relativo a convenções sociais. Depois lidamos com o utilitarismo, a teoria ética segundo a qual uma ação é moralmente correta a depender de suas consequências, e comparamos a visão utilitarista com a ideia deontológica de que uma ação é moralmente correta se for motivada pelo dever. Por fim, discutimos sobre a relação entre filosofia moral e filosofia política e examinamos algumas tentativas de entender o que significa dizer que uma sociedade é *justa*.

Antes de avançar, eu preciso enfatizar que este é um livro concebido para ser lido com calma, com pausas. Não hesite em fazer anotações, sublinhar, rabiscar (ou o que você preferir), indicando o que lhe parece mais importante, porque um livro sem marcas é um livro não lido.

1
Filosofia, senso comum e ciência

Neste capítulo, nós exploramos uma resposta minimamente consensual à pergunta sobre o que é a filosofia, o que também envolve tratar de algumas questões adjacentes, como, por exemplo, questões sobre a natureza de problemas filosóficos. Note que, se as pessoas fazem filosofia ainda hoje, é porque existem problemas, alguns dos quais sem dúvidas são genuínos, que elas querem resolver — e elas os resolvem (ou assim pretendem fazer) com respostas e explicações de certo tipo. Portanto, em alguma medida vamos falar de respostas e explicações filosóficas, além de problemas filosóficos. Antes de começar, no entanto, vamos fazer uma limpeza de terreno.

Antes de qualquer coisa...

Talvez você já tenha ouvido falar que a filosofia é algo como a "busca pela verdade absoluta". Essa concepção é tentadora, eu admito, mas ela está errada — e por mais de uma razão. Em primeiro lugar, não faz sentido falar em "verdade absoluta". Uma qualificação como 'absoluta' sugere que existe uma oposição entre a suposta verdade absoluta e uma verdade "pela metade", uma verdade meia-boca, por assim dizer. Talvez o que as pessoas tenham em mente quando dizem algo do tipo seja mais ou menos isto: "a filosofia almeja uma verdade para além das *minhas* verdades e das *suas* verdades, e por isso ela é uma busca pela verdade absoluta". Aí que está o problema: a "sua verdade" ou a "minha verdade" são expressões que não querem dizer

nada mais, respectivamente, do que aquilo em que você acredita e aquilo em que eu acredito. Pense por um momento em alguma coisa qualquer em que você acredita. Por exemplo, que você está diante de um livro. Se você acredita nisso, é porque acredita que seja verdade. Essa seria a "sua verdade", para usar as palavras do nosso interlocutor imaginário. Mas o problema com esse vocabulário — que eu sugiro erradicar daqui em diante — é que ele sugere que aquilo em que você acredita seria uma "verdade relativa". Não é: o que você acredita é somente isso, aquilo em que você acredita. E você pode inclusive estar errado, sua crença pode ser *falsa*. Portanto, não é a "sua verdade" que você tem um livro diante de si, é nada mais do que uma crença ou opinião formada de um modo ou de outro. Por essa razão, isto é, porque não faz sentido falar em "verdades relativas" ou "verdades subjetivas", é que não faz sentido falar em "verdades absolutas" (do mesmo modo, não faz sentido falar em "fatos alternativos", porque um fato alternativo simplesmente não é um fato, é uma interpretação dos fatos). E por isso a filosofia não é uma busca pela verdade absoluta.

Em segundo lugar, a filosofia não pode ser apenas uma busca pela verdade. Bom, o que exatamente é a verdade é assunto do nosso quarto capítulo. Antes de tentar lançar luz sobre esse conceito, podemos descartar que a filosofia seja entendida apenas como a busca pela verdade porque essa caracterização é abrangente demais. De fato, filósofos buscam a verdade. O problema é que outras disciplinas também a buscam. Quando uma bióloga se pergunta sobre a estrutura de um organismo, ela está buscando a verdade. Quando uma astrônoma se pergunta sobre o que causa a perturbação na órbita de um planeta, ela está buscando a verdade. Assim como quando uma química se pergunta por que uma reação ocorre. Aliás, quando eu me pergunto pelo que haverá de almoço, ou quando você se pergunta quantas páginas faltam até o próximo capítulo, nós também estamos buscando a verdade. A busca pela verdade, portanto, não é exclusiva da filosofia, e por isso não pode ser o que a define (ou seja, filosofia não pode ser *só* buscar a verdade).

Dito isso, parece que a filosofia possui algo em comum com as ciências, assim como pode ter algo em comum com certas atitudes de dia-a-dia. Com efeito, em todas as dimensões da nossa vida, nós nos deparamos com certas dificuldades, com certas perplexidades. Diante dessas perplexidades, nós tentamos dar sentido à nossa experiência através de *explicações*. Ou seja, nós tentamos explicar a nossa experiência, para nós e para outros, e com isso procuramos entendê-la. Isso que queremos dizer quando falamos que uma disciplina como a filosofia (ou como a biologia, astronomia, etc.) "busca a verdade": ela quer fazer sentido, oferecer entendimento, sobre certas coisas.

Vamos começar com nossas explicações de dia-a-dia e ver como elas funcionam. Essas explicações são tipicamente bastante simples e podem ser fruto da nossa experiência imediata ou herdadas da nossa comunidade. Além disso, frequentemente não possuem muitas aberturas para revisões. Essa é uma caracterização, ainda que muito superficial, de como formamos nossas crenças de senso comum (sem tom pejorativo aqui). Por exemplo, temos uma explicação de senso comum de que, quando o ar está muito quente e abafado, é porque vai chover. Isso é algo que nós podemos atestar, seja pela nossa experiência individual, seja pelo relato de outras pessoas em quem confiamos. Ao fazermos isso, formamos uma crença de que vai chover porque o ar está muito quente e abafado. No nosso dia-a-dia, a não ser que tenhamos algum interesse mais refinado em meteorologia, não dispomos de razões muito sofisticadas para explicar esse fenômeno (dia quente antes da chuva) de modo mais aprofundado, tampouco para colocar aquela explicação em dúvida. Basta que as coisas assim nos apareçam, e seguimos adiante.

Também as ciências empíricas — isto é, ciências que dependem de experimentação —, podem explicar um fenômeno como a relação entre o calor e a umidade relativa do ar. Podem inclusive mostrar que a explicação de senso comum é inadequada. Não é, portanto, o *objetivo* de uma crença ou de uma explicação que distingue ciência de senso comum, porque ambas podem querer explicar a mesma

coisa, mas sim o *modo* como essa explicação é estabelecida. Uma explicação científica é mais sofisticada do que uma explicação de senso comum, porque envolve certos procedimentos que a comunidade científica reconhece como confiáveis, envolve certa *técnica*, por assim dizer. Além disso, uma explicação científica caracteriza-se por uma atitude *crítica* ou *refletida*. Isso quer dizer que, na ciência, o que é dito geralmente depende de razões para ser aceito e está aberto para revisão (por outros cientistas competentes). Com efeito, é plausível que as incertezas sejam o principal motor da prática científica: sem que dúvidas razoáveis sejam levantadas, não há progresso na ciência. Em alguns casos, até repostas muito bem consolidadas podem ser revistas a depender de novas evidências. Veremos isso em detalhes no nosso quinto capítulo. Por enquanto, basta entender que a ciência busca a verdade de modo diferente do senso comum. De um modo muito geral, podemos dizer que ambas iniciativas tentam produzir entendimento sobre a nossa experiência, mas explicações científicas são *críticas* e *técnicas*, enquanto explicações de senso comum, podemos dizer, geralmente são irrefletidas e simples (não que isso faça com que elas sejam *erradas*!).

E quanto à filosofia, será que ela também se encarrega de oferecer explicações como o senso comum ou como a ciência?

Comecemos pela relação entre senso comum e filosofia. Dissemos acima que, diante de certas perplexidades, buscamos explicações. Essas perplexidades podem ser simples, corriqueiras. Você pode se perguntar, por exemplo: por que o dia ficou tão tenso, quente e carregado, mas não choveu? Às vezes, você pode se deixar levar e ir um pouco mais além: qual será a relação entre a umidade do ar e a chuva? Será que um *causa* o outro? E se você tem tempo em mãos, pode ir ainda mais longe nos seus questionamentos: *o que significa dizer que uma coisa causa a outra*? Ou ainda: *por que do fato de que eu observei isso acontecer assim-e-assado no passado eu espero que aconteça do mesmo modo no futuro*? Essas últimas perguntas, eu sugiro, já entram no reino das perplexidades propriamente filosóficas. Uma vez que são perguntas que estão no plano de fundo das nossas práticas ordi-

nárias e que, por conta disso, não podem ser respondidas do mesmo modo que as demais, considero-as *perplexidades fundamentais*.

Responder as nossas perplexidades fundamentais requer certo tipo de atenção e de cuidado de que geralmente não dispomos no decorrer do nosso cotidiano — mas já que você está aqui, vou assumir que você tenha um tempinho. É o que nós vamos fazer neste livro: voltarmo-nos a questões e problemas que podem ocorrer naturalmente no nosso dia-a-dia, embora nem sempre nós nos detenhamos nesses assuntos. Você com certeza já passou várias vezes por algum questionamento desse tipo e, portanto, você não é estranho à prática filosófica, ainda que talvez desconheça o vocabulário específico com que fazemos filosofia. Você já deve ter se perguntado, por exemplo, ao ver alguém cometer uma atrocidade, *como essa pessoa pode considerar isso certo*? Ou: *devo fazer isso porque todo mundo faz a mesma coisa*? Ou ainda, ao ouvir alguém dizer que sabe de algo tremendamente implausível, contra todas as evidências: *como essa pessoa poderia saber disso?* Todas essas perguntas podem facilmente ser conduzidas a uma investigação genuinamente filosófica e, por causa disso, indagações filosóficas estão em continuidade com nossas perplexidades cotidianas. Desse modo, a busca pela verdade na filosofia não é radicalmente diferente da busca pela verdade no dia-a-dia (ou, o que equivale, de acordo com nossas opiniões e explicações de senso comum).

Mas a filosofia, diferentemente do senso comum, não é irrefletida. Na prática — de acordo com o modo como filósofos genuínos fazem filosofia legítima, não como impostores cometem seus simulacros de pensamento — respostas filosóficas devem sempre poder ser justificadas. Ou seja, um pronunciamento filosófico não é como uma *revelação* de alguém que se apresenta como um guru. Um pronunciamento filosófico deve ser acompanhado de boas razões, e não de meras especulações conspiratórias ou devaneios febris sem quaisquer fundamentos. Isso também significa que não há posicionamento filosófico que *não possa* ser questionado. Claro, nós não podemos contestar *tudo* de uma única vez, pois temos de manter fixos alguns resultados e algumas respostas enquanto avançamos — mas, a cada

novo passo, podemos voltar atrás e contestar o que havíamos dado por certo até então (naturalmente, desde que tenhamos boas razões para tal).

Além disso, explicações filosóficas também envolvem procedimentos que a comunidade filosófica reconhece como adequados, envolvem certa técnica cujo domínio não pode prescindir da prática (como tudo que envolve técnica). Como nas ciências, a técnica filosófica é exibida pela proficiência em um vocabulário, isto é, por saber usar a linguagem específica da disciplina. A linguagem filosófica frequentemente envolve termos homônimos aos que usamos no nosso dia-a-dia, mas cujo significado no domínio filosófico é completamente diferente do significado com que os empregamos cotidianamente. Esse fato, aliás, serve-nos de advertência: em primeiro lugar, devemos estar sempre atentos a termos técnicos que podem sugerir certa familiaridade, porque isso pode levar ao erro (um bom texto filosófico sempre mantém claro o significado de seus termos — a obscuridade é o expediente de um escritor que não sabe sobre o que está pensando). Em segundo lugar, por mais que tentemos evitar termos técnicos neste livro, eles são de fato inevitáveis. Sem eles, nosso trabalho seria ao mesmo tempo hercúleo e superficial. Termos técnicos em excesso, contudo, tornariam as nossas discussões inacessíveis a um leitor ou a uma leitora de primeira viagem. Por essas razões, vamos procurar manter um equilíbrio entre as linguagens cotidiana e técnica no decorrer deste livro. Podemos perder um pouco de profundidade com isso, mas ganhamos em acessibilidade.

Seria, então, a investigação filosófica como a científica? De acordo com os critérios de criticidade e de tecnicidade, parece que sim. Mas isso não é tudo. Com efeito, tudo que vimos até aqui é amplamente consensual, mas daqui em diante há divergências profundas.

Duas concepções de filosofia

A pergunta sobre a qual vamos nos ocupar agora é se a filosofia é como a ciência. Assim posta, no entanto, essa é uma pergunta vaga.

Em certo sentido, é indisputável que a filosofia "é como a ciência", porque ambas atendem (ou almejam atender) os critérios de criticidade e de tecnicidade. Mas parece que há uma diferença fundamental entre o modo como é conduzida a investigação filosófica e o modo característico da investigação científica. Considere, por exemplo, como uma cientista procura estabelecer um resultado particular, digamos, na biologia. Imagine uma bióloga que procura mostrar que os organismos X e Y têm um ancestral comum, Z. Para fazer isso, ela deve recorrer a ferramentas e procedimentos típicos da pesquisa biológica. Por exemplo, observar semelhanças estruturais, genéticas, etc., entre X e Y, especular sobre as pressões ambientas que podem ter resultado na seleção de certos traços em detrimento de outros e sobre como eram os ambientes habitados pelos ancestrais de cada organismo, e assim por diante. Tudo isso requer o que chamamos acima de *experimentação* ou *experiência*. Requer — com o perdão da metáfora — confrontar-se com o mundo. A bióloga não pode chegar ao resultado que almeja, isto é, mostrar que o organismo X e o organismo Y têm um ancestral comum Z, *apenas* pensando a respeito. Ela deve confrontar-se com o mundo através de observações, deve realizar experiências. Ou seja, embora obviamente o pensamento seja essencial para a investigação da bióloga, sem experiência, sua investigação não passaria de uma especulação vazia.

Agora, se dermos um passo adiante na ordem de abstração, um passo em direção às nossas perplexidades fundamentais, podemos nos perguntar o seguinte: *o que diferencia um ser vivo de um ser não vivo?* Essa pergunta não é como a pergunta da bióloga, e talvez não possamos respondê-la do mesmo modo. Não parece que conduzir uma observação, ou um conjunto de observações, vá nos ajudar muito aqui. Podemos inclusive observar vários seres vivos (pessoas, gatos, cachorros, pintassilgos) e listar o que esses seres têm em comum, e ao mesmo tempo observar vários seres não-vivos (pedras, cadeiras, televisores) e listar o que esses seres têm em comum. Vamos supor que uma lista desse tipo *possa* ser feita — o que parece implausível de qualquer maneira. Mas, supondo que seja possível, poderíamos dizer

que todo ser vivo tem a característica *A*, e todo ser não-vivo não tem essa mesma característica. Mesmo que tivéssemos uma lista desse tipo, nós ainda não explicaríamos o que diferencia um ser vivo de um ser não vivo. A pergunta filosoficamente interessante passaria a ser: *por que a característica* A *faz com que seres vivos sejam vivos*? Ou ainda, *por que a ausência de* A *faz com que seres não-vivos não sejam vivos*? O ponto aqui é que, assim como perplexidades filosóficas podem surgir de questões do nosso dia-a-dia, elas também podem surgir a partir de investigações científicas. De acordo com isso, quando cientistas se propõem a responder a esse tipo de questão, eles estariam fazendo filosofia (mesmo que talvez não saibam).

De qualquer modo, uma resposta satisfatória a essas questões não pode ser dada pela experiência, alguns diriam, porque em última análise é uma pergunta sobre o nosso *conceito* de *ser vivo*. Ou seja, não é uma pergunta primariamente sobre o mundo, sobre os seres vivos eles mesmos, mas sobre o que significa, para nós, pensar que algo é um ser vivo, concebê-lo como tal. Responder a essa pergunta envolveria uma *análise* ou um *esclarecimento conceitual*, algo que não é feito empiricamente, mas a partir de reflexões, raciocínios e especulações. Por exemplo, será o caso que nós consideramos um organismo qualquer como vivo porque ele faz coisas, vai a lugares, interage com seu ambiente? Um aspirador de pó automático também se comporta dessa maneira, mas não estaríamos dispostos a dizer que ele é vivo. Talvez o nosso conceito de *ser vivo* faça referência aos conceitos de *alma* ou de *espírito* — mas daí alguém poderia argumentar que podemos imaginar um ser, digamos, uma inteligência artificial muito sofisticada, capaz de comportar-se exatamente como nós, mas que hesitaríamos ao dizer que ela tem um espírito. Outros poderiam argumentar que não haveria por que negar que uma inteligência artificial *tão* sofisticada tenha um espírito (alma, mente, etc.), e que ela seria mais como seres humanos e menos como aspiradores de pó automáticos — esse é aproximadamente o argumento principal do filme *Blade Runner* e da série *West World*, entre tantos outros. Ou também poderíamos dizer: será que um ser é vivo porque tem liberdade

para agir, isto é, porque não está preso a regras da natureza? Mas será que *nós* temos esse tipo de liberdade?

Talvez um exemplo melhor ainda sobre a diferença metodológica entre filosofia e ciência seja o famoso *problema do mal*. O problema do mal é geralmente apresentado como uma prova da inexistência de Deus. Uma versão muito simples desse argumento seria a seguinte: vamos supor, conforme a caracterização usual, que Deus seja bom, onisciente, onipotente e onipresente. É um fato incontestável que o mal (sofrimento, miséria, etc.) existe no mundo. Como pode Deus coexistir com o mal? Se Deus permite a existência do mal, então Deus não é bom. Se Deus não sabe da existência mal, então não é onisciente. Se Deus não pode intervir e evitar que o mal aconteça, então não é onipotente. E se Deus não está presente onde há o mal, então Deus não é onipresente (se está presente onde há o mal, então não é bom). Portanto, argumentariam os ateus, Deus não existe. Se esse problema for encarado do ponto de vista filosófico, e não apenas como um problema a ser resolvido pela fé, então cabe aos teístas mostrar o que poderia estar errado com pelo menos um dos passos que levam à conclusão de que Deus não existe. Diferentemente das investigações científicas, no entanto, teístas não podem responder àquele argumento com base em observações empíricas. Não é como se pudessem sair por aí procurando por Deus embaixo de mesas, atrás de portas ou no topo de montanhas, pois não é isso que solucionaria o problema do mal. Talvez o caminho seja interpretar o que queremos dizer com os conceitos de *mal, bom, onisciente,* etc. Isso, no entanto, é radicalmente diferente de fazer ciência, é análise ou esclarecimento conceitual.

Segundo a concepção tradicional de filosofia, portanto, a filosofia não é uma investigação sobre o mundo, ela é uma investigação sobre os nossos conceitos, os modos através dos quais nós podemos pensar o mundo. Está claro que essa ideia pressupõe que apenas podemos acessar o mundo através de conceitos, o que poderia ser contestado, mas isso não vem ao caso aqui. Em virtude de tratar-se primariamente de conceitos, e apenas secundariamente da realidade, a concepção

tradicional de filosofia entende a filosofia como uma *disciplina de segunda ordem*. Ou seja, as disciplinas científicas são disciplinas de *primeira ordem*, elas tratam do mundo ele mesmo. A filosofia, de acordo com isso, não trata do mundo ele mesmo, ela trata dos nossos pensamentos sobre o mundo. Daí então a ideia de que a filosofia e as ciências são radicalmente distintas—ainda que, eventualmente, no curso da prática científica, questões filosóficas apareçam. Respondê-las, no entanto, não envolveria o mesmo método que responder questões empíricas. Envolveria pensar sobre os nossos conceitos. Alguns filósofos, na primeira metade do século XX, deram um passo adiante e afirmaram que, como todos os nossos conceitos são expressos em uma linguagem, a filosofia consiste em *análise de linguagem.*

Embora a ideia de filosofia como disciplina de segunda ordem seja persuasiva, ela não é a única na disputa. Essa caracterização depende da suposição de que conceitos não são parte do mundo, mas uma espécie de reino separado da realidade. Mas por que, alguns filósofos perguntaram, deveríamos conceber conceitos como não sendo parte da realidade, do mundo? Se, por outro lado, conceitos são parte do mundo, assim como é a linguagem através da qual nós expressamos esses conceitos, e se a filosofia consiste em análise ou esclarecimento conceitual, então a filosofia trata do mundo ele mesmo, não menos do que as disciplinas científicas. Essa linha de pensamento constitui uma vertente do chamado *naturalismo filosófico*, a ideia de que a filosofia é contínua com as ciências naturais.

'Contínua' ainda não é uma expressão muito clara. Vimos que mesmo a ideia de filosofia como disciplina de segunda ordem aceita uma continuidade fraca entre ciência e filosofia, pois admite que, apesar de ambas serem críticas e técnicas, existe uma diferença radical de método entre elas. O que os naturalistas querem dizer com 'continuidade' é algo muito mais forte, a saber, que a filosofia investiga o mundo assim como a ciência. Ou seja, quando uma filósofa se pergunta, por exemplo: *o que é a ação?* Ou: *qual a coisa certa a se fazer?* Ou ainda: *que tipo de coisas existem?* Ela não está apenas se perguntando sobre como nós pensamos nos conceitos de *ação, moralidade,*

existência. Ela está se perguntando sobre coisas no mundo, e nada impede que a resposta dela seja informada pelos avanços científicos.

Dentro da ideia naturalista mais geral, podemos distinguir três correntes: o naturalismo *ontológico,* o naturalismo *epistemológico* e o naturalismo *metodológico*. O nome do primeiro vem da palavra grega para 'existência'. Segundo essa versão, só podemos assumir como existente aquilo que as ciências naturais dizem que existe, a saber, entidades espaço-temporais. O segundo, cujo nome também vem do grego, dessa vez da palavra que remete ao conhecimento, diz que o único conhecimento legítimo é aquele obtido pelos métodos científicos. Isso, é claro, não parece razoável, visto que eu posso saber que a gata Ostara está sobre a mesa sem conduzir nenhuma investigação científica (eu descubro isso meramente olhando para ela). Por fim, o terceiro tipo de naturalismo alega que o método filosófico é o mesmo que o método da ciência, e que consiste em experimentação e observação. No caso da filosofia, observamos como usamos certos conceitos e com isso formulamos hipóteses. Depois, submetemos essas hipóteses a testes através de experimentos. Assim como nas ciências, em que uma hipótese pode ser refutada (ou refinada) por um experimento empírico que demonstre a sua inadequação, uma hipótese filosófica pode ser refutada (ou refinada) por um experimento *mental*, a imaginação de um cenário possível, que demonstre sua inadequação. Nesses casos, consultamos o que nós diríamos se as coisas fossem assim-e-assado, ou seja, consultamos nossas disposições para julgar o que nos é apresentado e com isso avaliamos a plausibilidade de uma teoria. O que é importante é que, segundo o naturalista, fazemos tudo isso sem sair do reino da realidade, sem comprometer-se com a ideia de que a filosofia é uma disciplina de segunda ordem. No terceiro capítulo, nós veremos em detalhes que tipo de considerações motivam a ideia de que há uma continuidade robusta entre filosofia e ciência.

Considerações finais

O que afinal de contas é a filosofia? Em alguma medida, nosso esforço até aqui foi de mostrar o que a filosofia *não* é. Mas nós também vimos que problemas filosóficos podem nascer naturalmente das nossas perplexidades cotidianas ou prosaicas — com a diferença de que as perplexidades características da filosofia não parecem poder ser resolvidas do mesmo modo que as nossas questões de dia-a-dia. Diferentemente das nossas perplexidades cotidianas, que são resolvidas por explicações e crenças de senso comum, a filosofia é mais bem caracterizada como uma disciplina crítica ou refletida, no sentido de que seus enunciados devem poder ser justificados através de razões e que, além disso, enunciados filosóficos não são imunes à dúvida. É sempre possível perguntar por que um pronunciamento filosófico deve ser aceito — se não for possível levantar essa pergunta, não estamos mais no campo da filosofia, mas do charlatanismo. O aspecto crítico da filosofia é especialmente importante porque mostra que fazer filosofia requer *coragem intelectual*, por assim dizer, a coragem de perguntar pelas razões por trás de certas respostas. Além disso, problemas filosóficos às vezes também aparecem no horizonte da prática científica, que também é crítica ou refletida, no mesmo sentido que a prática filosófica o é. Ambas as disciplinas são altamente técnicas, o que significa que exigem o manejo de um vocabulário especializado, com neologismos ou termos homônimos aos que usamos no dia-a-dia. Mas, ainda que haja amplo consenso sobre essa taxonomia inicial, há um ponto de divergência: seria a filosofia uma disciplina que investiga o mundo ele mesmo, como as demais ciências, ou uma disciplina de segunda ordem, uma espécie de pensar sobre o pensamento?

Antes de examinar as razões em favor de cada posição, talvez seja importante esclarecer uma última questão. Talvez você encontre por aí alguns acadêmicos que dedicam todas as suas forças para *defender* um autor ou uma autora e para interpretar exaustivamente seus escritos, dissecando todas as suas vírgulas. Às vezes ocorre, sobretudo

no meio acadêmico brasileiro, que alguns estudiosos mantenham seu “filósofo de estimação”, uma espécie de santo no altar, de modo que tudo que contrarie a palavra desse filósofo estaria errado ou não faria sentido. Isso é profundamente adverso ao espírito crítico e investigativo da filosofia. Por isso eu defendo que a interpretação de outros filósofos, a história da filosofia, por mais que seja um *meio* crucial e indubitavelmente interessante para a investigação filosófica, jamais deveria ser a *finalidade* da filosofia. A filosofia almeja produzir respostas às nossas perplexidades fundamentais—alguns diriam que isso ocorre na forma de esclarecimentos conceituais, outros diriam que isso ocorre na forma de resolução de problemas reais, tão reais quanto os que ocupam as cientistas. Deixando essa disputa de lado por um momento, o que importa é que ambos os lados concordam sobre o seguinte ponto: a filosofia visa produzir entendimento, seja sobre os nossos conceitos, seja sobre o mundo ele mesmo. Desse modo, responder às nossas inquietações, por exemplo, sobre o que é ter conhecimento, ou sobre o que é moralmente correto, com enunciados do tipo “Fulano diz que é conhecimento é isso-e-aquilo” e “Beltrano diz que a moralidade é isso-e-aquilo” — mesmo que Fulano e Beltrano sejam notórios pensadores de barba longa, toga ou peruca — seria insatisfatório. A não ser que a resposta continue da seguinte maneira: “..., mas estão errados porque isso-e-aquilo”, ou, o que é mais raro: “... e estão corretos porque isso-e-aquilo”, limitar-se a uma exegese de um respeitável pensador não atende às nossas perplexidades iniciais. Nós queremos saber o que é, nesses exemplos, conhecimento e moralidade, não o que algum eminente filósofo disse que é.

Se você não está convencido com meu argumento acima, cuja conclusão é de que a mera interpretação do que disseram outros filósofos não deve ser a finalidade da filosofia (ainda que possa ser um valioso meio); eu o(a) convido a abrir um livro de qualquer grande filósofo que seja comumente elevado à santidade pelos seus exegetas (você conhece alguns: gregos barbudos, franceses excêntricos, ou alemães com nomes com muitas consoantes). Não importa quem você porventura escolher, você verá que o filósofo em questão não

está *apenas* interpretando outros filósofos que vieram antes de si. Por uma razão simples: um filósofo digno do nome também quer oferecer respostas para problemas filosóficos genuínos — e, ainda que para isso tenha que interpretar seus interlocutores e discordar de seus predecessores, o que de fato é muito frequente, certamente sua filosofia não se limita a isso. Devemos seguir o exemplo de grandes pensadores, discutir suas respostas, mas sempre com a coragem de acrescentarmos a nossa avaliação sobre o sucesso ou o fracasso das suas investidas intelectuais.

Leituras recomendadas

Este livro se beneficia da consulta da *Enciclopédia de termos lógico-filosóficos*, organizada por Branquinho, Murcho e Gomes (2006). A introdução e o primeiro capítulo de Williamson (2019) são um excelente ponto de partida—mas eu insisto que você leia o livro todo. O livro introdutório escrito por Russell (2008) explora outros problemas filosóficos e é considerado uma obra clássica da filosofia contemporânea. Para um entendimento profundo sobre o que caracteriza a filosofia analítica, o livro de Glock (2011) é indispensável. Para uma reconstrução do argumento do problema do mal, eu recomendo Bruce e Barbone (2013). O nosso enfoque não é histórico, mas, se esse for seu interesse, você encontra observações interessantes em Jordan (2016) sobre o surgimento da filosofia.

2
Algumas distinções fundamentais

Vimos no capítulo anterior que uma concepção de filosofia a distingue das ciências por concebê-la como uma disciplina de segunda ordem. Isso significa, de acordo com essa concepção mais tradicional, que a filosofia é uma espécie de pensamento sobre o pensamento. Mas, é claro, essa ainda é uma caracterização vaga, até porque outras disciplinas, como a psicologia, a antropologia e a sociologia, também estudam o pensamento humano sob perspectivas variadas. Naturalmente, a concepção tradicional de filosofia a diferencia daquelas ciências. Para mostrar como isso é feito pelos proponentes dessa ideia, será necessário familiarizarmo-nos com alguns termos técnicos e distinções importantes, a saber, *a priori* x *a posteriori*, *analítico* x *sintético* e *necessário* x *contingente*.

A priori e *a posteriori*

Inquéritos filosóficos, nós vimos, parecem independer da experiência. Pense novamente no problema do mal, que mencionamos no capítulo anterior. Um teísta que pretende refutar a conclusão daquele argumento — a saber, que Deus não existe — não pode recorrer à experiência, ou seja, simplesmente tentar *observar* a existência de Deus. Ele tem de recorrer a outros meios que não envolvam a experiência, como a análise conceitual dos termos envolvidos naquele argumento.

O mesmo acontece em outros debates filosóficos. Por exemplo: há uma familiar discussão em ética, que se originou na antiguidade clássica e que ressurgiu na modernidade, sobre se a moralidade é relativa ou absoluta. Algumas pessoas argumentam que a moralidade é relativa, seja a cada um de nós, ou a cada cultura, com base no suposto fato de que há grande variação de critérios morais de pessoa para pessoa ou de cultura para cultura. Esse suposto fato, sem dúvidas, é fruto de uma observação empírica — não por acaso esse tipo de concepção relativista da moral tornou-se popular com o avanço da antropologia a partir do século XVIII, que passou a registrar as diferenças entre os costumes dos povos europeus e os dos povos ameríndios ou africanos. Mas, apesar da motivação empírica, notemos que o mero fato de que existe uma variação nas convenções sociais desses povos não garante que a moral seja relativa. Com efeito, é perfeitamente possível (um absolutista moral argumentaria) que a moral seja objetiva e que exista *apenas um* conjunto de regras morais, não obstante a suposta variação nas práticas através de culturas. Ocorreria, se o absolutista estiver correto, que algumas culturas podem estar erradas, algumas culturas podem estar certas, ou até mesmo todas erradas simultaneamente. De qualquer modo, tanto a posição relativista quanto a posição absolutista não podem ser justificadas com base na experiência. Até porque elas não se pretendem descrições de como as pessoas *de fato* agem, elas dizem respeito à natureza de conceitos morais. Para tomar uma decisão a respeito, não é suficiente nem necessário *observar* e *descrever* como as pessoas se comportam, de cultura para cultura, mas sim lançar luz sobre os nossos *conceitos* envolvidos em juízos morais. Isso não é uma empreitada empírica, é uma investigação de natureza diferente.

A distinção que nos interessa aqui, portanto, é entre o que é *dependente da experiência* e o que é *independente da experiência*. A concepção de filosofia como disciplina de segunda ordem afirma que a filosofia é independente da experiência, enquanto as ciências são dependentes da experiência. Mas a que exatamente esses termos se aplicam?

Essa distinção foi popularizada pelo filósofo alemão Immanuel Kant no século XVIII, e a partir do seu trabalho passamos a usar as famosas expressões *a priori* e *a posteriori* para referirmo-nos, respectivamente, ao que é independente da experiência e ao que é dependente da experiência. Na filosofia kantiana, a distinção aplica-se a *juízos*, de tal modo que um juízo é *a priori* se ele independe da experiência, e *a posteriori* se ele depende da experiência. Não é tão óbvio, contudo, que um juízo *a priori* independa da experiência (em algum sentido). Primeiro: não parece claro que eu preciso da experiência para adquirir o conceito de *computador*? Parece-me que sim. Vamos supor que o conceito de *computador* seja algo como: *máquina que processa informações e que apresenta resultados de acordo com normas simbólicas de operação*. Se eu nunca tivesse sido apresentado a um computador, se eu tivesse vivido na época de Kant, por exemplo, eu nunca teria tido o conceito de computador. Portanto, a experiência é indispensável para a aquisição desse conceito. Mas eu posso pensar: 'um computador é um computador'. Não é um juízo muito útil, não é um juízo que vá me ajudar muito no meu dia-a-dia, mas — e isso é interessante — é um juízo cuja verdade eu posso atestar *sem depender da experiência*. Sem investigar o mundo, eu posso vir a saber que é verdade que um computador é um computador. Ou seja, do ponto de vista da sua *verificação* — da descoberta da verdade desse juízo — ele é *a priori*, ainda que o conceito que eu emprego nesse juízo tenha sido adquirido por meio da experiência.

Mas, é claro, a filosofia kantiana não é exatamente o estado da arte já faz alguns séculos. Hoje sabemos que 'juízo' é um termo ambíguo: ele pode referir-se a um *conteúdo do que é pensado* — por exemplo, o que é expresso pelo enunciado 'um computador é um computador' — ou ao produto de uma atividade intelectual, digamos, *a justificação para crer* ou o *conhecimento* de que um computador é um computador. Enunciados (ou, como usarei intercambiavelmente aqui, frases) por si não são *a priori* nem *a posteriori*, eles são, como veremos, analíticos ou sintéticos, e sua verdade (se verdadeiros) é necessária ou contingente. Por enquanto, não importa tanto precisar o que queremos

dizer com a expressão 'enunciado', basta ter em mente que um enunciado é um tipo de uso da linguagem que visa descrever um estado de coisas. Proferir as palavras 'a porta é branca', com o significado que geralmente atribuímos a elas, é proferir um enunciado. Um enunciado pode ser verdadeiro ou falso, em virtude de tentar descrever alguma coisa — se bem-sucedido, é um enunciado verdadeiro, e é falso caso contrário. Diferentemente de 'a porta é branca', 'me passa a manteiga' não é um enunciado, porque não descreve um estado de coisas, mas expressa uma ordem. Portanto, não pode ser verdadeiro ou falso. O mesmo vale para questões, como 'a porta é branca?'. Uma questão não é verdadeira ou falsa, mas ela admite uma resposta, essa sim que pode ser verdadeira ou falsa.

De qualquer modo, para evitar confusão, trataremos a distinção entre *a priori* e *a posteriori* como uma distinção a respeito de tipos de *conhecimento* e evitaremos a expressão kantiana de "juízos *a priori*/*a posteriori*". No capítulo seguinte, veremos como essa escolha pode ser revista, e que talvez um defensor da distinção entre *a priori* e *a posteriori* deva dizer que ela se trata da *justificação para crer*. Antes disso, é importante ressaltar que, dada essa caracterização dos pares da distinção como um modo de obtenção de conhecimento, segue-se que o modo como ela frequentemente é usada fora da filosofia envolve um erro grosseiro, a saber, como algo que acontece *antes* ou *depois* da experiência. A distinção, como vimos, não trata de eventos sob uma perspectiva temporal, portanto é inadequado usá-la daquela maneira. Do mesmo modo, às vezes se usa equivocamente a expressão '*a priori*' com o significado de 'a princípio', mas é claro que pode haver princípios que são justificados *a posteriori*.

O conhecimento filosófico, então, se é possível de ser obtido, é obtido *a priori* (segundo a concepção tradicional de filosofia). O conhecimento científico, por oposição, é obtido *a posteriori*. Sabemos que no caso deste último isso envolve a experiência, a observação, a confrontação com o mundo, em um sentido mais geral. Mas o que exatamente envolve a aquisição de um conhecimento *a priori*?

Conhecimentos *a priori* parecem ser típicos, por exemplo, da matemática. Eu não preciso saber nada sobre o mundo para saber que todo triângulo tem três lados. Há, nesse enunciado, uma generalidade envolvida ("*todo* triângulo...") que impede que a descoberta sobre a verdade desse enunciado possa ser feita empiricamente. Quando falamos de *todo* triângulo, estamos falando de todas as figuras geométricas compostas apenas de três ângulos, desde o passado imemorial até o futuro inatingível. Como poderíamos acessar a verdade daquele enunciado empiricamente? Como poderíamos observar *todo* triângulo? A resposta, é claro, é que não podemos. E, no entanto, é indubitável que todo triângulo tem três lados, nós sabemos disso. Parece, pois, que somos capazes de acessar a verdade desse enunciado através de uma faculdade puramente mental e não-experimental, como que se pudéssemos *ver* essa verdade com uma máxima segurança, sem risco de erro. Vamos chamar essa faculdade de acesso *a priori* à verdade de certos enunciados de *intuição racional*.

Como quase tudo na filosofia, contudo, a ideia de uma faculdade cognitiva que permite acessar (ou simplesmente intuir) *a priori* a verdade de certas proposições não é consensual. Muitos críticos da ideia de intuição racional apontaram que essa é uma faculdade misteriosa, que não *explica* o conhecimento *a priori* porque o próprio funcionamento dessa capacidade não está explicado. E talvez sequer seja possível fazê-lo, porque, para um proponente da intuição racional, explicar *a priori* a faculdade cognitiva que gera conhecimentos *a priori* requereria o exercício dessa própria faculdade. Se, no entanto, for possível explicar *a priori* como funciona a intuição racional sem depender dela, é porque talvez ela não seja o principal modo de aquisição de conhecimento *a priori*. É tentando evitar esse tipo de problema que muitos filósofos buscaram explicar o conhecimento *a priori* a partir de frases analíticas, que caracterizaremos a seguir. Segundo essa concepção, o conhecimento *a priori* tem sua peculiaridade porque é sobre enunciados especiais, como 'todo triângulo tem três lados' — enunciados cuja verdade não depende de como o mundo é.

Antes disso, resta avaliar outra boa candidata ao título de fonte de conhecimento a priori, a saber, a reflexão. Embora pareça plausível que podemos obter conhecimento a priori através da reflexão, aqui há um problema. Refletir, pensar sobre os próprios pensamentos, é claramente uma atividade que transcorre no tempo e que envolve recursos como atenção, concentração e memória. Isso não é nenhuma novidade para as ciências cognitivas. Além disso, quando nos confrontamos com o mundo, também precisamos de atenção, concentração e memória para torná-lo inteligível a nós. Dadas essas semelhanças, portanto, não haveria por que assumir que a reflexão é substancialmente diferente da investigação empírica. Por outro lado, se entendermos que o que divide o conhecimento empírico do conhecimento *a priori* é a operação dos sentidos (visão, audição, etc.), e que a reflexão não é mediada por sentido algum, como se fosse apenas o voltar da atenção à nossa própria mente e às suas operações; então um conhecimento obtido através da reflexão seria *a priori*. No momento, não nos interessa fazer uma escolha aqui, apenas ter em mente que há uma possibilidade de divergência quanto ao que conta como *a priori*.

Analítico e sintético

A priori e *a posteriori*, nós vimos, dizem respeito ao modo como o conhecimento é adquirido. Ou seja, trata-se de uma distinção *epistêmica*. Já enunciados são entidades linguísticas, e a eles é aplicada outra distinção. Considere estes exemplos:

1. A bola é verde.
2. Nada que é totalmente verde é totalmente azul.
3. Alguns dentistas são ricos.
4. Um ortodontista é um dentista.
5. Getúlio Vargas está morto.
6. Se Getúlio Vargas se matou, então Getúlio Vargas está morto.
7. Maria é formada em direito.
8. Se Maria é promotora de justiça, Maria é formada em direito.

Você nota alguma diferença entre as frases pares e ímpares? Especialmente no que diz respeito à verdade delas? Note, por enquanto, que não estamos preocupados em definir o que é a verdade, isso é nosso trabalho para o quarto capítulo. Por enquanto, tenha em mente uma concepção incontroversa sobre verdade: um enunciado é verdadeiro se é bem-sucedido em descrever como as coisas são.

Se você pensou cuidadosamente, pode ter notado que parece haver uma distinção com respeito à verdade das frases pares em relação às frases ímpares. Como a outra distinção, essa também foi popularizada por Kant. Para entendê-la, considere a frase 3., 'alguns dentistas são ricos'. Se a separamos em sujeito e predicado, teremos 'alguns dentistas' como sujeito, e 'são ricos' como predicado. Importante: essa separação é um tanto antiquada do ponto de vista da lógica contemporânea, mas como Kant tinha essa separação em mente (ou algo próximo a ela) ao distinguir analítico e sintético, ela vai nos servir por enquanto. Se examinarmos o conceito expresso pelo sujeito do enunciado em 3., 'alguns dentistas', não vamos encontrar o conceito de *rico*, que é expresso pelo predicado. O conceito de *dentista* é tão somente o conceito de uma pessoa com uma formação específica, que trabalha de determinada maneira. Ou seja, esse conceito não diz nada sobre a riqueza ou a posse dessa pessoa. Por isso, prosseguiria Kant, a verdade desse tipo de enunciado é como uma *síntese*, pois nele são combinados elementos que não estão originalmente relacionados nos conceitos expressos pelo sujeito e pelo predicado. Por essa razão, os enunciados ímpares no nosso exemplo são *sintéticos*.

Compare agora com o enunciado 4., a saber: 'um ortodontista é um dentista'. Se fizermos o mesmo procedimento de quebrar o enunciado em partes, veremos que o conceito de *ortodontista* inclui ou contém o conceito de *dentista*. Justamente, uma ortodontista é uma dentista que trabalha com deformidades dos dentes e suas correções. Desse modo, a verdade desse enunciado não envolve a combinação de conceitos que estão separados, mas, pelo contrário, envolve analisar o conceito de ortodontista em partes mais simples. Negar esse enunciado levaria a uma contradição — 'nem todo ortodentista é um

dentista' —, o que não é o caso para enunciados sintéticos, pois não é contraditório dizer que nem todo dentista é rico. Por essas razões, os enunciados pares no nosso exemplo acima são chamados de *analíticos*. Parece haver, portanto, uma diferença substancial entre a verdade de enunciados analíticos e a verdade de enunciados sintéticos.

Mas há (pelo menos) um problema grave com a distinção traçada nos moldes kantianos. Se você observar os enunciados 2, 6 e 8, verá que eles não têm a forma de sujeito e predicado (como em 'a porta é branca'), mas sim de uma frase condicional do tipo *se... então*. Não faz sentido, à primeira vista, falar em conceito do sujeito conter o conceito do predicado nesses casos, o que limitaria a aplicação da distinção kantiana. É necessária, portanto, uma nova estratégia para explorar a distinção, que até agora parece bastante plausível, entre enunciados sintéticos e analíticos.

Gottlob Frege, outro filósofo alemão, ofereceu uma reformulação promissora da distinção kantiana no final do século XIX. Assim como Kant, Frege acreditava na validade da distinção, mas, contra Kant, Frege defendia que os enunciados da aritmética, como '2 + 2 = 4', são *analíticos*, ao passo que Kant defendia que enunciados como esse são *sintéticos*. O projeto magistralmente inaugurado por Frege consistia mostrar como enunciados da aritmética podem ser reduzidos a frases da lógica e, para isso, ele teve de reinventar a lógica, que havia basicamente estagnado desde as remotas descobertas de Aristóteles. Em lógica e filosofia da linguagem, Frege foi um visionário e um gênio, o precursor do que hoje chamamos de Filosofia Analítica (cujo nome não tem exatamente a ver com a ideia de enunciados analíticos), que foi o maior movimento filosófico do século XX e que ainda hoje é um movimento mundialmente robusto. Voltando ao assunto em questão, a ideia de Frege é aproximadamente a seguinte: um enunciado é analítico somente se pode ser convertido em um enunciado da lógica a partir da substituição dos seus termos por termos sinônimos e por definições. Considere, por exemplo, 6: 'se Getúlio Vargas se matou, então Getúlio Vargas está morto'. Podemos substituir 'se matou' por um sinônimo, digamos, 'está morto por conta própria'. Temos 6* (as-

terisco aqui denota um novo enunciado): 'se Getúlio Vargas está morto por conta própria, Getúlio Vargas está morto'. Se abstrairmos que esse enunciado trata do antigo presidente e ditador Getúlio Vargas, e usarmos um termo qualquer como '*g*' para denotar esse sujeito, temos: 'se *g* está morto por conta própria, então *g* está morto'. Agora, se separarmos o predicado complexo 'está morto por conta própria' em dois predicados, 'está morto' e 'por conta própria', e substituí-los por '*F*' e '*G*', respectivamente, temos: 'se *g* é *F* e *g* é *G*, então *g* é *F*'. Esse é um enunciado cuja verdade é uma questão de lógica: se algo qualquer tem duas propriedades, de ser *F* e de ser *G*, então esse algo tem a propriedade de ser *F*. Na perspectiva fregeana, portanto, um enunciado é analítico se é redutível a um enunciado da lógica através da substituição por sinônimos e definições, e sintético caso não seja possível convertê-lo em um enunciado da lógica.

Mais tarde, quando a filosofia analítica já estava a pleno vapor, o grupo de filósofos alemães e austríacos conhecidos como os empiristas (ou positivistas) lógicos do Círculo de Viena, cuja filosofia vigorou desde o final de 1920 até 1950, operava com outra distinção de analítico e sintético. O grupo era constituído por filósofos e físicos com interesses comuns que se reuniam a partir da iniciativa do alemão Moritz Schlick, que então ocupava a cátedra de Filosofia das Ciências Indutivas na Universidade de Viena. Para esses filósofos, um enunciado é analítico se a sua verdade depende do significado das palavras *apenas*. Por exemplo: o enunciado 'se Maria é promotora de justiça, Maria é formada em direito' é tal que o significado das palavras que o compõem é suficiente para determinar que é verdadeiro. Já o enunciado 'Maria é formada em direito' não é verdadeiro *apenas* em virtude do significado das palavras que o compõem, e sim em virtude do significado das palavras *e* do modo como o mundo é constituído, a saber, de Maria ter obtido uma titulação em direito.

Há, portanto, essas três maneiras de traçar a distinção entre analítico e sintético. Será essa distinção plausível? Ela parece correta à primeira vista. Mas veja que existem algumas dificuldades envolvendo essa distinção, a maioria delas dizendo respeito à analiticidade (a

propriedade de um enunciado ser analítico), como veremos no capítulo seguinte. Uma dificuldade é o famoso *paradoxo da análise*: se o conceito do sujeito já contém o conceito do predicado (ou, se há uma relação de sinonímia em enunciados analíticos, ou se a verdade deles depende apenas do significado das palavras), então como pode um enunciado analítico ser *informativo*? Para algumas pessoas, entender que um ortodontista é um dentista que trabalha com deformidades nos dentes e suas correções pode ser esclarecedor, trazer uma descoberta minimamente substancial sobre o mundo. Se, no entanto, não há uma relação de contenção, ou uma relação de sinonímia, etc., em enunciados analíticos, o que por sua vez explicaria o seu caráter informativo, então como poderiam ser analíticos? Como eu disse, essa é apenas uma dificuldade, e veremos mais delas no capítulo seguinte. Mas antes disso, e antes de entender como essa distinção auxilia na nossa compreensão de filosofia como uma disciplina de segunda ordem, vamos trabalhar com ainda outra distinção bastante plausível.

Necessário e contingente

Certas verdades parecem ser diferentes de outras. Considere o seguinte enunciado: 'Maria é promotora de justiça'. Vamos considerar que seja um enunciado verdadeiro, isto é, que exista uma pessoa chamada Maria e que ela trabalhe como promotora de justiça. Agora, poderia ser o caso que Maria nunca tivesse existido (talvez aí caberia dizer que o enunciado não faria sentido — mas, de qualquer modo, não seria verdadeiro), tanto como poderia ser o caso que Maria trabalhasse com qualquer outra coisa ou que estivesse aposentada. Qualquer uma dessas possibilidades mostra que a verdade do enunciado 'Maria é promotora de justiça' é *contingente*. Ou melhor: é um enunciado contingentemente verdadeiro — há cenários alternativos em que ele seria falso. Chamamos esses cenários, seguindo o filósofo alemão Gottfried Leibniz, de *mundos possíveis*.

Uma questão polêmica na filosofia contemporânea diz respeito à natureza de mundos possíveis: seriam eles mundos que existem pre-

sentemente, como que se fossem realidades alternativas? Ou seriam mundos imaginados ou, mais precisamente, estipulados a partir da nossa (a única) realidade? Qualquer que seja a resposta aqui, por enquanto basta entender que um dos mundos possíveis é o mundo *atual*, o mundo em que vivemos hoje. Um enunciado contingentemente verdadeiro, portanto, é tal que, em primeiro lugar, é verdadeiro no mundo atual e, em segundo lugar, há pelo menos um mundo possível em que ele é falso. Note que há enunciados que são *contingentemente falsos*: um enunciado é contingentemente falso se é falso no mundo atual, mas há pelo menos um mundo possível em que é verdadeiro, por exemplo: 'agora está chovendo'. Vamos supor que não esteja. Mas o ponto é que *poderia* estar (o tempo verbal aqui é crucial).

Contraste 'Maria é promotora de justiça' com o seguinte enunciado: 'Maria é Maria'. Claro, não é um enunciado tremendamente esclarecedor — veja, no entanto, que, se o nome denota alguém, ou seja, se há alguém chamada 'Maria', esse enunciado não apenas é verdadeiro, mas é *necessariamente* verdadeiro. A ideia aqui é que todos os mundos possíveis em que o nome designa Maria, esse é um enunciado verdadeiro. Se fosse um enunciado em que o nome não designa ninguém, como em 'Belsnickel é Belsnickel', alguns o considerariam vacuamente verdadeiro (porque aquilo que é inexistente é igual a si mesmo), enquanto outros o considerariam um enunciado sem significado, apenas que parece ser significativo. Não precisamos posicionarmo-nos sobre isso por enquanto. Agora, um enunciado como 'chove *e* não chove', ao referir-se ao mesmo local em um mesmo tempo, não apenas é falso, como *necessariamente* falso. Ou seja, não há mundo possível em que é verdadeiro que chove e não chove. Mesmo em Salvador da Bahia, em que a chuva é um fenômeno tão imprevisível e efêmero, *ou chove ou não chove*, mas nunca *chove e não chove*.

A distinção entre necessidade e contingência é uma distinção modal, porque nesse caso diz respeito ao modo de verdade de enunciados. Pode ficar mais claro visualizar essa distinção da seguinte maneira:

	Necessário	Contingente
Verdadeiro	Verdadeiro em todos os mundos possíveis	Verdadeiro no mundo atual, falso em pelo menos um mundo possível
Falso	Falso em todos os mundos possíveis	Falso no mundo atual, verdadeiro em pelo menos um mundo possível

Além de verdades e falsidades necessárias e contingentes, nós também falamos de *possibilidades*. Um enunciado contingentemente verdadeiro é um enunciado possivelmente verdadeiro — ou, de modo mais direto: se é verdade no mundo atual, é possível que seja verdade. Pense no caso de 'Maria é promotora de justiça'. Se isso é verdade, ainda que seja contingentemente verdadeiro, é possível que seja verdade. Note que um enunciado contingentemente verdadeiro também é possivelmente falso. De modo semelhante, um enunciado contingentemente falso também é possivelmente verdadeiro, como é o caso do enunciado: 'Klaus é um gato preto'. Klaus atualmente é branco, mas é perfeitamente possível que haja um mundo possível próximo em que Klaus teria sido um gato preto. Agora, se pensarmos, por exemplo, em algo que é necessariamente falso, temos uma impossibilidade, como '2 + 2 = 5'. Não é possível que esse enunciado seja verdadeiro, porque ele é falso em todos os mundos possíveis. Por fim, um enunciado necessariamente verdadeiro é *possivelmente* verdadeiro também. Se é uma verdade necessária que Maria é Maria, é possível que Maria seja Maria.

Nós caracterizamos acima um enunciado como um uso da linguagem que almeja descrever a realidade e que, portanto, pode ser verdadeiro *ou* falso. Ou seja, a tentativa pode ser bem-sucedida, mas também pode falhar. Se eu dissesse, diante de uma parede branca, que ela é roxa (ou verde ou azul, etc.), eu teria proferido um enunciado *falso*, porque eu teria falhado na tentativa de capturar a realidade com aquele enunciado. Mesmo assim, teria sido um enunciado legítimo, ainda que fosse falso. Mas, note bem, se um enunciado necessariamente verdadeiro é tal que ele *não poderia ser falso* — isto é, que

não há mundo possível em que ele é falso — será que se trata de um enunciado legítimo? O filósofo austríaco Ludwig Wittgenstein, um pensador tão brilhante e excêntrico quanto sua obra é tortuosa e impenetrável, foi um dos autores que desenvolveu essa dúvida (não exatamente com essas palavras) no começo do século XX. Segundo ele (aparentemente...!), a filosofia tradicionalmente almejou oferecer enunciados que, se verdadeiros, seriam necessariamente verdadeiros. Como, no entanto, esses enunciados não são legítimos porque (trivialmente) não podem ser falsos, seriam apenas *pseudo*-enunciados, ou melhor, afirmações sem sentido. A filosofia não faria sentido, conclui Wittgenstein — mas, para mostrar esse ponto, ele teve que fazer uma quantidade substancial de filosofia. O empenho filosófico seria, na metáfora da qual se apropriou, como uma escada que você joga fora depois de ascender através dela. Depois de supostamente ter resolvido todos os problemas da filosofia no seu único livro publicado em vida, o *Tractatus Logico-Philosophicus*, de apenas 75 páginas, Wittgenstein abandonou a filosofia, tornou-se jardineiro num monastério, depois professor primário num vilarejo, eventualmente ameaçou um eminente filósofo com um atiçador de lareira... e depois se deu conta de que estava errado, de que não havia resolvido todos os problemas da filosofia.

As distinções na prática filosófica

Vimos três importantes distinções, e agora estamos em condições de situá-las em uma grande taxonomia filosófica geral, o que nos ajudará a delinear com mais clareza a ideia de filosofia como disciplina de segunda ordem.

Intuitivamente, parece que as classes de enunciados sintéticos, de enunciados cujas verdades são contingentes e de conhecimentos obtidos *a posteriori* são classes coextensivas. Ou seja: um enunciado como ‘Maria é promotora de justiça’ é um enunciado sintético, a verdade é claramente contingente e podemos ter conhecimento acerca do fato de que Maria é promotora de justiça apenas empiri-

camente. Do mesmo modo, parece claro que enunciados analíticos, enunciados cujas verdades são necessárias e conhecimentos obtidos *a priori* são classes também coextensivas. O enunciado 'Maria é Maria' é analítico (a depender de como nós interpretamos a noção de analiticidade — o modo como Kant traça a distinção claramente não se aplica aqui, porque não é um enunciado composto de sujeito e predicado, apesar das aparências, e sim um enunciado de identidade), é necessariamente verdadeiro e pode ser descoberto sem nenhuma experiência.

Com efeito, esse é o modo como os empiristas lógicos concebiam a relação entre essas três distinções. Em certos aspectos, esses filósofos seguiam a tradição dos empiristas clássicos da modernidade, uma corrente que foi especialmente forte no mundo britânico. A linha central do empirismo é que todo o conhecimento é, como o nome sugere, oriundo da experiência empírica. De acordo com os empiristas lógicos, então, o conhecimento científico é oriundo da experiência empírica, isto é, adquirido *a posteriori*. Além disso, os enunciados científicos são sintéticos, pois são sobre o mundo, e as descobertas científicas são contingentemente verdadeiras, porque dizem respeito ao modo como o mundo é, e não a uma relação entre conceitos. Para os empiristas lógicos, apenas enunciados da lógica e da matemática (se verdadeiros) são necessariamente verdadeiros, bem como demais enunciados envolvendo definições. A filosofia, afirmavam esses filósofos seguindo Wittgenstein, não deveria propor enunciados substanciais sobre a realidade, mas — e esta é uma contribuição original do empirismo lógico — *reconstruir logicamente* (*a priori*) a linguagem científica através de enunciados que eliminavam obscuridades e compromissos filosoficamente escusos. Essa concepção é, portanto, concordante com a ideia de que a filosofia é uma disciplina de segunda ordem, claramente distinta das ciências empíricas.

O que melhor diferencia os empiristas lógicos dos seus predecessores modernos é, além de um refinamento lógico sem precedentes, o *critério verificacionista do significado*. Esse critério, a principal linha de pensamento compartilhada pelos membros do Círculo de Viena

durante boa parte da sua duração, afirma que o significado de um enunciado consiste nas suas condições de verificação. Ou seja, o que um enunciado qualquer significa, como, por exemplo, 'a parede é branca' consiste nas condições sob as quais é possível verificar empiricamente que a parede é branca. Desse modo, um enunciado *sintético conhecido a priori* seria para eles uma aberração — pois, de acordo com o modo como eles traçam a distinção entre analítico e sintético, um enunciado é sintético se a sua verdade depende de como é o mundo. Como seria possível verificar *a priori* (isto é, independentemente da experiência) um enunciado cuja verdade depende de como o mundo é? Isso seria claramente absurdo. Por essa razão, enunciados sintéticos conhecidos *a priori não fazem sentido* de acordo com esses filósofos.

A ideia de que enunciados sobre o mundo (isto é, sintéticos) podem ser conhecidos *a priori* é a própria ideia da famosa — ou famigerada, dependendo da sua perspectiva — *metafísica*. No discurso popular, a expressão 'metafísica' ficou indevidamente associada a algo místico, oculto, mas o termo técnico filosófico não tem nada a ver com isso. Embora o nome provavelmente tenha sido cunhado porque os tratados de Aristóteles que lidavam acerca do "ser enquanto ser" haviam sido dispostos logo *após* os seus tratados sobre física — daí a ideia de que seriam livros "que vêm depois da física", em grego simplesmente *metafísica* — o termo acabou colando. Com efeito, a concepção de Aristóteles da "filosofia primeira" ou da "ciência do ser enquanto ser" é bem apropriada para mostrar com o que lida a metafísica. Podemos estudar o ser, ou melhor, as coisas que existem (para usar um vocabulário menos arcaico) de várias maneiras. Podemos estudá-las da perspectiva da biologia, da física, ou ainda da perspectiva das suas relações sociais. Para usar a expressão aristotélica, essas seriam todas ciências do ser *enquanto* alguma outra coisa: enquanto ser biológico, ou ser físico, ou ser social e assim por diante. Mas quando estudamos o ser *enquanto ser*, *enquanto ele próprio*, entramos no campo da metafísica. Quando nos perguntamos, por exemplo, que tipo de coisas existem ou se há uma relação causal entre elas,

estamos fazendo metafísica. Plausivelmente, essa disciplina consiste em um conhecimento básico, elementar, razão pela qual Aristóteles escolheu chamá-la de *filosofia primeira*. Esse nome, no entanto, foi substituído pelo título menos pomposo de *metafísica*, e hoje também a chamamos de *ontologia*.

Retomando, então: os empiristas lógicos buscavam exonerar a metafísica da filosofia porque a consideravam sem sentido (dado que os enunciados metafísicos são sintéticos e pretensamente conhecidos *a priori*). Mas os empiristas lógicos não são nem de longe a maioria na história da filosofia. Outras escolas filosóficas defenderam a possibilidade do conhecimento *a priori* de enunciados sintéticos, isto é, a possibilidade da metafísica. Notoriamente, a tradição que nasceu na modernidade com o nome de *racionalismo* (que é diferente do que às vezes se chama de racionalismo hoje na linguagem popular!) defendia que é possível obter conhecimento substancial acerca do mundo sem depender da experiência. Os racionalistas modernos não usavam essas palavras — pois eles em grande maioria precederam Kant, e o vocabulário contemporâneo é herdado deste último. Em geral, racionalistas não negavam que possamos obter conhecimentos do mundo através dos sentidos, portanto, dependentemente da experiência, e que os enunciados em questão seriam sintéticos. Seu ponto, simplificadamente, é que nosso conhecimento do mundo não pode prescindir de conhecimentos *a priori*.

Talvez o melhor exemplo disso seja a filosofia de Descartes. Filósofo, físico e matemático francês do século XVII, Descartes foi um proeminente racionalista. Nas suas famosas *Meditações Metafísicas*, o personagem através do qual Descartes escreve argumenta que, ao notar que mantivera em algum momento opiniões falsas, as suas opiniões presentes também poderiam ser falsas. Diante da possibilidade de engano, seria impossível erigir — para usar a metáfora arquitetônica característica do texto — o edifício do conhecimento humano. Ou seja, Descartes assume que a possibilidade de erro é incompatível com o conhecimento, tese que hoje chamamos de *infalibilismo*. Portanto, seria preciso fornecer fundamentos seguros e inabaláveis

ou, mais precisamente, crenças não apenas indubitadas, mas *indubitáveis* para que se candidatassem ao título de conhecimento. Para chegar a essas crenças indubitáveis, ele avalia a fonte de todos os supostos conhecimentos obtidos até então, a percepção. Tudo que aparentemente sabemos, Descartes argumenta na *Primeira Meditação*, parece dar-se através dos sentidos, da percepção sensorial. Portanto, para eliminar as opiniões que não sejam indubitáveis, basta lançar dúvida sobre as fontes do nosso (suposto) conhecimento, os sentidos.

Ele prossegue com um sofisticado, elusivo e sutilmente retórico exercício de imaginar uma dúvida cada vez mais radical e abrangente. O que sobreviver (se algo sobreviver) à versão mais radical da dúvida será a base a partir da qual deverá ser possível fundamentar o conhecimento sem chances de engano. Seria suficiente, para levantar essa dúvida, constatar a falha ocasional dos sentidos? Sem dúvidas meus sentidos já me enganaram no passado, especialmente quando eu não estava em condições ideais de observação, como quando eu confundi na rua uma pessoa qualquer com um conhecido meu. Mas e quando eu me encontro em condições ideais de observação, digamos, concentrado, lúcido, sóbrio e diante de um objeto à curta ou à média distância? Como eu poderia estar errado, por exemplo, a respeito da minha impressão de ver uma gata preta sentada na mesa à minha frente? Eu estou atento, ciente, ainda não bebi hoje (a quarentena de 2020 mudou radicalmente nossos hábitos). Como poderia estar errado?

Descartes observa que eu *poderia* estar sonhando que vejo uma gata preta na mesa à minha frente — e que, o que é pior, enquanto sonho, eu tenho a mesma convicção de que o que eu percebo é real como quando estou acordado e lúcido, de fato percebendo uma gata. Ou seja, da perspectiva da primeira pessoa, sonho e percepção verídica são indistinguíveis. Portanto, não é possível eliminar, do ponto de vista da primeira pessoa, a possibilidade de estar sonhando. O que mostra que mesmo num caso ideal de observação, eu poderia estar enganado.

Isso é plausível, mas insuficiente para mostrar o ponto que ele pretende: poderia haver conhecimentos que escapam da hipótese do sonho. Descartes fala em conhecimentos que envolvem elementos como números e cores, que não poderiam ser sonhados, enquanto apenas a combinação desses elementos poderia ocorrer em um sonho. Talvez seja mais claro pensarmos em enunciados da matemática: embora o sonho possa afetar a nossa percepção sensorial, ele não é capaz de afetar o conhecimento de que 2+3=5, pois este suposto conhecimento possui outra fonte, plausivelmente a nossa intuição racional (para Descartes, ao menos). Para levantar a dúvida sobre um conhecimento tão claro e distinto como um conhecimento matemático, Descartes introduz a figura do Deus Enganador, que logo é substituída — por razões não completamente transparentes no texto — pelo Gênio Maligno. A possibilidade do Gênio Maligno consiste em supor a existência de um ente infinitamente poderoso cujo único intento é causar o tormento epistemológico no protagonista das meditações. O Gênio Maligno, assim, faria com que qualquer suposto conhecimento obtido, seja por qual meio que for, esteja falso. É importante notar que Descartes não afirma que o Gênio Maligno existe, mas sim que ele *poderia* existir e que, *se* existisse, não haveria como eliminar essa possibilidade de engano promovida por ele. Ou seja: tudo que nós aparentamos saber está suspenso dada a possibilidade de engano massivo gerada pelo Gênio Maligno. Descartes finalmente chegou na sua dúvida hiperbólica, uma dúvida que mostra que todas as opiniões, mesmo aquelas que ele acreditava serem seguras e invioláveis, podem ser falsas.

Ao final da *Primeira Meditação*, então, perdeu-se tudo. Tudo que se parecia saber, todas as opiniões e convicções mantidas até então *podem* ser falsas — e isso, para um infalibilsita, é suficiente para derrogar o suposto conhecimento adquirido a partir dessas opiniões. Para voltar ao nosso assunto principal, notemos como Descartes pretende livrar-se dessa enrascada metafísica a partir da *Segunda Meditação*. Sua estratégia, em linhas gerais, consiste em 3 passos. O primeiro passo é notar que, mesmo que ele esteja sistematicamente enganado,

há algo acerca do qual ele não pode enganar-se, a saber, que ele está pensando. 'Pensamento', para Descartes, é um termo técnico que abrange as atitudes de crer, duvidar, especular, lembrar, raciocinar, enganar-se etc. Desse modo, mesmo que eu mantivesse uma opinião falsa, enganosa, a respeito do mundo em virtude da ação do Gênio Maligno — vamos imaginar que seja a crença de que a gata preta está sobre a mesa — eu posso ter uma crença a respeito do que eu mesmo acredito: *eu penso que a gata preta está sobre a mesa* (por ser uma crença sobre o que eu acredito, chama-se crença de *segunda ordem*). A crença de que a gata preta está sobre a mesa pode muito bem ser falsa; a crença de que eu creio que a gata preta está sobre a mesa, não! A descoberta brilhante de Descartes é que, na primeira pessoa do singular, do tempo presente indicativo, eu tenho acesso infalível, indubitável aos meus pensamentos. Esse é o primeiro passo que ele apresenta para reconquistar o mundo perdido.

O segundo passo é mostrar que, se eu penso, eu existo. É o famoso *Cogito Cartesiano*: eu penso, eu existo. Note bem, isso não é a mesma coisa que dizer que *tudo que pensa existe, eu penso, logo, eu existo* — não, esse não é o seu ponto, e isso seria um erro grosseiro. O ponto de Descartes é que há algo pensante mesmo quando eu me engano. A esse algo, Descartes atribui o nome de "substância pensante" (em oposição à substância corpórea, que é espaço-temporal). Isso é motivo de mais controvérsias do que podemos ocupar-nos aqui. De qualquer modo, até esse ponto, então, Descartes provou que eu penso (mesmo que esteja enganado sobre o que estou pensando, não posso enganar-me sobre o fato de estar pensando) e que eu existo — claro, não eu, o autor deste livro nem você que o lê, tampouco o próprio Descartes, mas sim *um eu qualquer* que passa pelo procedimento imaginado por Descartes.

Mas, mesmo assim, isso ainda é muito pouco: e as outras pessoas, os animais, as árvores, a praia, e até mesmo o meu passado? Se tudo que Descartes provou é que eu penso, eu existo, então estou preso no exíguo espaço *solipsista*, um espaço habitado apenas pela minha própria mente (o que não é muito diferente do ponto de vista prático da

quarentena por causa do novo coronavírus). Daí vem a mais curiosa e ambiciosa estratégia de Descartes: provar que não estou sozinho, porque estou na companhia de Deus (o bom, não o enganador). Mas como Descartes prova que Deus existe?

Em primeiro lugar, ele não apresenta essa prova apenas uma única vez, de um único modo. Isso, aliás, deve deixar-nos alerta: se a prova fosse boa, não seria necessário repeti-la várias vezes e sempre de modo diferente. Em segundo lugar, na sua versão mais clara, ela é extremamente simples, pois parte de apenas duas premissas que são, à primeira vista, bastante razoáveis e incontroversas: primeiro, que Deus possui todas as perfeições. Segundo, que existir necessariamente é uma perfeição. Com isso Descartes conclui que Deus existe necessariamente.

Parece fácil. A primeira premissa quer dizer que Deus é perfeito: sumamente bom, onisciente, onipresente, etc. Podemos aceitar isso sem maiores dificuldades. Com a segunda premissa Descartes quer dizer que uma existência contingente, como a nossa, que possui começo, meio e fim, e que aliás poderia nem ocorrer, é uma existência *imperfeita*. Uma existência perfeita, por oposição, é uma existência necessária, inevitável. Algo que existe necessariamente existe desde sempre, para todo sempre, em todos os mundos possíveis. Logo, Deus, que possui todas as perfeições, existe necessariamente. Agora, solte um pouco este livro, pense por um momento e me diga: você está convencido pelo argumento de Descartes? Seja sincero. Se você não está, pode juntar-se à maioria dos filósofos, tanto os teístas quanto os ateístas. Talvez o próprio Descartes esteja incluso no grupo dos que não se convenceram completamente com essa que seria uma prova ao mesmo tempo tão simples e tão grandiosa (se bem-sucedida). Explicar o que está errado com essa prova é mais complicado, e só poderemos ver com clareza no quarto capítulo. Mas, eu peço que por enquanto você apenas confie em mim, algo está errado!

De todo modo, o que temos aqui é um caso paradigmático de uma (suposta) prova *a priori* de um enunciado sintético. Ou seja, Descartes quer mostrar que Deus existe, o que é um enunciado sintéti-

co, porque diz respeito ao mundo, ao tipo de coisa que existe. E ele faz isso com um procedimento completamente *a priori*. Mais do que isso: é porque sabemos que Deus existe, um Deus bom e não maligno, que podemos confiar nos nossos sentidos, visto que a existência de Deus garante que não podemos estar errados quando percebemos algo como claro e distinto. Ou seja, o bom funcionamento dos nossos sentidos como fonte de aquisição empírica de conhecimento é condicionado por uma prova *a priori* da existência de Deus. Essa prova de caráter metafísico é, pois, um caso paradigmático de racionalismo moderno.

Certamente ainda há racionalistas hoje, mas poucos seriam os que pretendem provar teses substanciais acerca da realidade, como a existência de Deus ou a imortalidade da alma, sobretudo depois do surgimento da filosofia kantiana. Um notório racionalista contemporâneo é Laurence BonJour, que pretendeu provar a indispensabilidade de recursos apriorísticos para a expansão do nosso conhecimento empírico. Seu argumento é simples e persuasivo. Considere as coisas que sabemos através dos sentidos. Por exemplo, eu sei (visualmente) que o gato branco está na sala e sei (através da memória) que a gata preta está no quarto. Mas como eu posso passar desses conhecimentos particulares para o conhecimento de que *há dois gatos nessa casa*? A mera percepção não me autoriza a realizar essa passagem, então algum outro tipo de recurso deve ser necessário, argumenta BonJour. Esse recurso seria uma intuição racional que me permitiria acessar a correção de um argumento da seguinte forma: *se o gato branco está na sala e a gata preta está no quarto, então há dois gatos nessa casa*. Será correto, no entanto, que, para expandir nossos conhecimentos empíricos, necessitamos de recursos apriorísticos? Uma resposta de teor empirista ao argumento de BonJour pode apelar aos fatos de que habitualmente fazemos inferências daquele tipo (entre muitas outras), e de que algumas são corretas, enquanto outras são incorretas. Eventualmente selecionamos as inferências que melhor se coadunam com as nossas estimativas do que seja uma inferência correta, ou revisamos as nossas estimativas de acordo com inferências que

são especialmente persuasivas. Isso dispensaria a necessidade de um recurso *a priori* para expandir nossos conhecimentos empíricos (voltaremos a esse assunto no quinto capítulo).

Como mencionamos, a filosofia original de Kant abalou fortemente as pretensões racionalistas, e com isso ele cunhou uma versão própria de filosofia como disciplina de segunda ordem. Kant, com efeito, é o pai dessa concepção, mas a maneira como ele a apresenta é peculiar. É virtualmente impossível fazer justiça à complexa e sofisticada filosofia kantiana em poucas linhas, então o que segue, você deve ter em mente, é uma simplificação quase grosseira, mas que basta para os nossos propósitos.

Na sua *Crítica da Razão Pura*, de 1781, Kant se pergunta como enunciados sintéticos conhecidos *a priori* são possíveis. Note: ele não se pergunta *se* eles são possíveis, eles claramente o são (para Kant). Ele afirma isso tendo por base que enunciados da matemática, embora sejam conhecidos *a priori*, não são analíticos, mas sintéticos. O enunciado 'cinco é igual a dois mais três' ('2+3=5') não é analítico, argumenta Kant, porque o conceito expresso pelo sujeito desse enunciado, a saber, *cinco*, não contém o conceito expresso pelo predicado, a saber, *ser igual a dois mais três*. Dizer que dois mais três é seis é um erro (um erro grosseiro até), mas não é uma contradição, como em 'chove e não chove'. Ou seja, trata-se de um enunciado sintético de acordo com o modo como Kant traça a distinção analítico/sintético, como vimos acima. (Aqui é o ponto de divergência entre Kant e Frege). Ao mesmo tempo, para descobrir a verdade desse enunciado, não é preciso empreender uma investigação empírica. Basta raciocinar, sem depender da experiência, que chegamos à verdade da frase de que cinco é igual a dois mais três. A pergunta, portanto, passa a ser: como são possíveis enunciados sintéticos *a priori*?

Pelo que dissemos acima, pode parecer que Kant é simpático ao projeto racionalista de fazer afirmações substanciais sobre o mundo de modo completamente independente da experiência. Mas Kant rechaça essas pretensões, porque o conhecimento sobre o mundo não pode prescindir da experiência — assim como não pode prescindir de

conceitos que organizam essa experiência de modo inteligível para nós. Ou seja, não é possível estabelecer verdades sintéticas *a priori* sobre o mundo ele mesmo. Para obter conhecimento, é preciso uma cooperação do que ele chama de faculdades da sensibilidade, que nos provém de representações do mundo, e do entendimento, que aplica conceitos a essas representações. Agora, para Kant, visto que todo o nosso acesso ao mundo ocorre por meio das representações obtidas através da nossa sensibilidade, articuladas conceitualmente pelo entendimento, não é possível ir para além das nossas representações e acessar como o mundo é independente de como nós o representamos. Ou seja, não podemos *transcender* as nossas condições sensíveis para acessar o mundo diretamente, como ele é, independente do modo como nós o representamos. *Mas* nós podemos investigar as nossas próprias faculdades cognitivas (isto é, que nos permitem acessar o mundo) para lançar luz sobre o modo como nós fazemos sentido do mundo. Por exemplo, a afirmação de que todo evento tem uma causa (veja, que não equivale à afirmação de que todo *efeito* tem uma causa, pois esta seria analítica) é um enunciado sintético que podemos conhecer *a priori* — mas é um enunciado cognoscível *a priori* porque diz respeito à *nossa própria maneira de interpretar e dar sentido* ao mundo, não ao mundo ele mesmo. Em resumo, essa é a maneira como Kant estabelece que a filosofia é uma disciplina de segunda ordem, uma disciplina distinta das ciências, à qual cabe investigar não o mundo, mas o modo como nos é possível acessá-lo.

Kant foi (e é) tremendamente influente, e não por acaso. Sua estratégia para lidar com problemas agudos da filosofia moderna é sofisticada e elegante, ainda que sua escrita tortuosa por vezes fique muito aquém disso, o que dificulta imensamente a compressão do seu trabalho e a apreciação da sua genialidade. Mas nada disso significa que Kant está correto. Em particular, a suposição de que todo o nosso acesso ao mundo é mediado por representações tem sido amplamente atacada, seja na filosofia da percepção, seja nas ciências cognitivas contemporâneas. Além disso, se Kant estiver correto sobre o fato de estarmos fechados, por assim dizer, pela nossa expe-

riência, como podemos afirmar que o mundo é a origem causal das nossas representações? Ou seja, afirmar que o mundo em si mesmo é inacessível não parece uma posição filosófica confortável, e talvez sequer seja coerente com a ideia de que o mundo de alguma maneira é responsável por causar as nossas experiências.

Considerações finais

Neste capítulo, procurei mostrar como as distinções entre analítico e sintético, *a priori* e *a posteriori* e necessário e contingente permitem lançar luz sobre a natureza da atividade filosófica segundo a concepção tradicional que delineamos anteriormente, a saber, a filosofia como uma disciplina de segunda ordem. Em especial é interessante notar que, embora a possibilidade de enunciados sintéticos conhecidos *a priori* tenha sido motivo de imensa controvérsia, ninguém havia questionado de modo contundente, até a segunda metade do século passado, os seguintes pontos: (1) a viabilidade da distinção entre analítico e sintético, o que consequentemente afetaria a viabilidade da distinção entre *a priori* e *a posteriori* e (2) as associações entre conhecimentos *a priori* e verdades necessárias e entre conhecimento *a posteriori* e verdades contingentes. Nós veremos esses desdobramentos no próximo capítulo.

Leituras recomendadas

Veja o oitavo capítulo de Russell (2008) para uma discussão sobre o *a priori*. Eu também recomendo enfaticamente os artigos de Teixeira (2014, 2015) para as questões da aprioricidade e da analiticidade. Casullo (2015) também oferece uma importante revisão das discussões clássicas e contemporâneas a respeito da aprioricidade, inclusive tratando de algumas objeções que veremos no próximo capítulo. Sobre necessidade, veja Motloch (2016). Você encontra passagens comentadas das *Meditações Metafísicas* de Descartes no segundo capítulo de BonJour e Baker (2010). Para

uma reconstrução da argumentação de Descartes, veja Hetherington (2013) e Lazier & Gaul (2013). Eu não recomendo ir direto à *Crítica da Razão Pura*, mas, se você insistir, use a tradução de Fernando Costa Mattos publicada como Kant (2018). Você encontra uma breve introdução à epistemologia kantiana em Silveira (2002) e uma defesa da acusação de que a epistemologia kantiana implica ceticismo em Williges (2007).

3

Novas perspectivas sobre antigas distinções

A ideia de filosofia como disciplina de segunda ordem implica conceber a filosofia como distinta da ciência. De acordo com isso, o conhecimento científico é *a posteriori*, os enunciados científicos são sintéticos, e a verdade desses enunciados (se verdadeiros, é claro) é contingente. A filosofia, ao contrário, seria uma fonte de conhecimento a priori, exibindo verdades necessárias. Na concepção tradicional, as divergências dizem respeito ao caráter analítico ou sintético dos enunciados filosóficos. Neste capítulo, exploraremos uma ideia alternativa, uma ideia que ataca as distinções que fundamentam a noção de filosofia como disciplina de segunda ordem. Além disso, veremos argumentos que, embora não sejam uma tentativa de motivar uma concepção alternativa de filosofia, acabam por mostrar que o modo aparentemente intuitivo como as distinções clássicas são tipicamente combinadas é problemático.

Analiticidade?

A distinção entre enunciados analíticos e sintéticos permaneceu amplamente incontestada desde sua canonização por Kant até a segunda metade do século XX. O jogo mudou em 1951, com a publicação de *Dois Dogmas do Empirismo* pelo filósofo norte-americano Willard van Quine. Com formação em matemática e inicialmente alinhado aos empiristas lógicos, que haviam migrado em peso para os Estados Unidos fugindo da ascensão nazista na Europa, Quine foi

uma das mais notáveis figuras da filosofia contemporânea, tornando-se um divisor de águas do movimento analítico — ainda que algumas de suas mais influentes ideias, como a continuidade entre filosofia e ciência, não fossem exatamente novas. Sete anos antes de publicar *Dois Dogmas*, Quine havia trabalhado no Brasil, na então Escola Livre de Sociologia e Política de São Paulo, em que ministrara aulas de lógica em português. Seu trabalho aqui deu origem a um livro, *O Sentido da Nova Lógica*, escrito originalmente na nossa língua com a ajuda do seu aluno, Vicente Ferreira da Silva, um precursor da lógica no Brasil que posteriormente se tornaria uma referência nacional em fenomenologia. Quine, segundo seu próprio relato, enturmou-se com ilustres figuras do cenário intelectual e artístico nacional, como Tarsila do Amaral, Mário de Andrade e Oswald Andrade. Não é uma imagem que você imaginaria facilmente: os modernistas de 22 e um filósofo analítico.

O impacto da publicação de *Dois Dogmas* foi tal que tornou a viabilidade da distinção entre analítico e sintético para sempre suspeita. Embora muitos filósofos tenham respondido a Quine *em defesa* da distinção, isso constituiu um esforço filosófico até então sem precedentes, o que serve de prova histórica, por assim dizer, de que a distinção não é tão óbvia quanto havia parecido. A ideia central no artigo é, como o nome sugere, explicitar o que havia sido até então aceito acriticamente, sobretudo pelos empiristas lógicos do Círculo de Viena. Uma tese aceita acriticamente é um *dogma*, que, sob inspeção mais cuidadosa, revela-se uma tese racionalmente infundada. O primeiro dogma criticamente avaliado por Quine é a distinção entre analítico e sintético, e o seu argumento consiste em mostrar que não há maneira de traçar uma distinção inteligível entre enunciados analíticos e sintéticos, porque não há como definir claramente, sem circularidade, o conceito de analiticidade. O segundo dogma é a ideia de que seria possível reduzir logicamente um enunciado teórico como ‘um elétron é uma partícula subatômica’ a um relato da experiência empírica. O ataque ao segundo dogma, por sua vez, acaba por mos-

trar que a distinção epistêmica entre *a priori* e *a posteriori* também está em julgamento.

O argumento contra o primeiro dogma, contra a distinção analítico e sintético, pode ser expresso do seguinte modo: nenhuma definição de analiticidade (pelo menos apresentada até então) é boa, porque todas elas fazem uso de conceitos que carecem de explicação ou até mesmo que recorrem à própria noção de analiticidade. Mais precisamente, lembremos que uma frase é considerada analítica se é *verdadeira em virtude do significado das palavras apenas*, como em 'todo solteiro é não-casado' (é assim que os empiristas lógicos definiam a noção de analiticidade, com quem Quine estava de fato discutindo nesse artigo). Supostamente a verdade dessa frase independe de como o mundo é, pois dependeria apenas do que significam as palavras contidas nela. Mas o que quer dizer essa definição? Como podemos explicá-la?

Uma ideia seria recorrer à noção fregeana de analiticidade. De acordo com essa ideia, uma frase analítica é tal que pode ser convertida em uma frase da lógica pela substituição de sinônimos por sinônimos, como podemos converter 'todo solteiro é não-casado' em 'todo não-casado é não-casado', porque 'solteiro' é sinônimo de 'não--casado'. Se esse for o caso, contudo, o defensor da analiticidade tem um problema. Note que tivemos de recorrer à noção de *sinonímia*, isto é, de palavras que possuem *o mesmo significado* para explicar o que significa dizer que uma frase é analítica. Ou seja, tentamos explicar como uma frase é verdadeira apenas em razão do significado usando indiretamente a noção de significado, que está pressuposta na noção de sinonímia. Portanto, ainda não explicamos o que é para uma frase ser analítica.

Quine continua dando voz a um interlocutor imaginário: talvez pudéssemos dizer que é possível entender o conceito de *sinonímia* através do conceito de *definição*. A definição de 'solteiro' é justamente 'homem não-casado', e isso pareceria explicar por que a substituição por uma expressão sinônima explica a verdade supostamente analítica de 'todo solteiro é não-casado'. Mas isso apenas nos diz quais

termos têm o mesmo significado que outros, sem, contudo, dizer o que afinal de contas é o conceito de *sinonímia*.

Talvez pudéssemos ainda dizer que duas expressões têm o mesmo significado, isto é, que são sinônimas, se for possível substituir uma pela outra em todas as ocasiões em que elas ocorrem sem prejuízo à verdade das frases em que ocorrem. Por exemplo, se 'Fulano é solteiro' é verdadeiro, ao substituir 'solteiro' por 'não-casado', não interferimos na verdade da frase inicial. Mas isso não nos explica ainda o que é, para duas expressões, terem o mesmo significado, porque algumas expressões são coextensivas — isto é, designam o mesmo conjunto de coisas — sem que tenham o mesmo significado. Pense, por exemplo, nas expressões 'seres humanos' e 'bípedes implumes'. O conjunto de seres humanos contém nós dois, eu suponho, além de muitas outras pessoas — exceto possivelmente alguns que abdicaram da sua humanidade ao defender projetos políticos hediondos, mas isso não vem ao caso. O problema é que esse conjunto contém os mesmos seres (todas as pessoas) que o conjunto de bípedes implumes. Dizer 'o autor deste livro é um ser humano' é verdadeiro, tudo leva a crer, assim como dizer que 'o autor deste livro é um bípede implume'. Mas obviamente 'ser humano' não tem o mesmo significado que 'bípede implume'!

Diante disso, talvez o que nos ajude seja o operador de *necessidade*, de modo que possamos dizer: os enunciados 'necessariamente, todo solteiro é solteiro' e 'necessariamente, todo solteiro é não-casado' são tais que permanecem verdadeiros se substituirmos 'solteiro' por 'não-casado'. Isso parece mostrar que 'necessariamente, todo solteiro é não-casado' é um enunciado analítico. O problema aqui, naturalmente, é com o advérbio 'necessariamente'. Porque, qualquer frase contendo o operador de necessidade seria considerada verdadeira justamente porque seria analítica. Ou seja, Quine aqui assume que uma frase necessariamente verdadeira o é em razão da sua analiticidade, independente de quaisquer peculiaridades que o mundo possa ter. Mas se isso for o caso, ainda não avançamos em definir o conceito

de *analiticidade* sem recorrer a outros conceitos que eles mesmos carecem de explicação, como o conceito de *necessidade*.

Como dissemos acima, a conclusão de Quine é pessimista com respeito à perspectiva da distinção: nenhuma definição do que queremos dizer com 'analiticidade' está livre de obscurantismos ou até mesmo de noções que dependem elas mesmas da noção de analiticidade. Ou seja: a noção é obscura ou circular e, portanto, não está bem explicada. Se não há uma boa explicação do que é para uma frase ser analítica, não há por que manter a distinção entre frases analíticas ou sintéticas. Talvez Quine esteja sendo excessivamente exigente com o que consideramos uma boa explicação, e esse é um ponto em que poderíamos pressioná-lo para defender a distinção. Além disso, pode-se argumentar que Quine focou numa versão *metafísica* da distinção (verdade em virtude do mundo/em virtude das palavras apenas), justamente porque tinha em mente os empiristas lógicos como interlocutores. Contudo, talvez o mais apropriado seja pensar uma versão *epistemológica* da distinção, segundo a qual uma frase é analítica quando o mero entendimento (que é uma noção epistemológica) dos seus termos é suficiente para apreender sua verdade. De qualquer modo, note que aquela distinção é crucial para diferenciar filosofia de ciência, porque, é claro, os enunciados científicos são tipicamente entendidos como sintéticos, e os enunciados filosóficos seriam, se verdadeiros, analíticos — sobretudo da perspectiva do empirismo lógico. Mas, se a distinção é espúria, seria um erro diferenciar filosofia de ciência! Esse é um primeiro e importante passo para estabelecer que a filosofia é contínua com a ciência quanto à natureza dos seus enunciados, o que caracteriza uma versão ampla de naturalismo filosófico.

Aprioricidade?

Com o segundo dogma, o problema ganha outra dimensão. Esse dogma diz respeito à possibilidade de reduzir logicamente um enunciado teórico (que seria sintético), como 'o elétron é uma partícula

subatômica' a um relato da experiência empírica. Essa ideia é crucial para os empiristas lógicos porque, como vimos no capítulo anterior, esses filósofos defendiam que o significado de um enunciado consiste nas suas condições de verificação — porque, afinal de contas, todo conhecimento tem origem da experiência empírica. O critério verificacionista do significado, por um lado, elimina a metafísica como uma disciplina cujos enunciados não têm significado — pois não é possível verificar a priori um enunciado sintético —, mas, por outro, faz com que alguns enunciados das ciências naturais se tornem perigosamente quase-metafísicos. Esses enunciados, por incluírem termos teóricos que se referem a entidades inobserváveis, parecem não fazer sentido porque não seriam verificáveis pela experiência empírica. Por exemplo, não é possível verificar que o elétron é uma partícula subatômica a partir da observação, porque 'elétron' é um termo que designa uma entidade que não somos capazes de observar. Para evitar problemas desse tipo, os empiristas lógicos recorrem à ideia de *redução*, segundo a qual um enunciado teórico seria redutível a um relato da experiência empírica, o que também explicaria como enunciados sintéticos são verdadeiros em virtude do mundo (diferentemente de enunciados analíticos). Assim, apesar de conter termos que se referem a entidades inobserváveis, esses enunciados teriam o mesmo conteúdo empírico que relatos experimentais, e não estariam em risco de serem confundidos com a famigerada e temível metafísica.

O problema dessa concepção torna-se evidente se considerarmos um cenário em que uma teoria científica qualquer (chamemo-la de T) é levada ao teste empírico. Vamos supor que T é confrontada por um conjunto de evidências E e, como frequentemente é o caso na prática científica, E não está de acordo com as nossas expectativas teóricas baseadas em T. Ou seja, o teste empírico e a teoria divergem, há alguma incongruência entre eles, pois as coisas não acontecem como esperávamos. Como, então, podemos situar o erro e corrigi-lo? Antes de responder, considere que testar uma teoria requer o uso correto de instrumentos científicos, a pressuposição de outras teorias mais

gerais de plano de fundo e até mesmo a descrição das evidências coletadas. Ou seja, no caso em que enunciados teóricos e evidências empíricas não se coadunam, o erro pode estar situado em diversos pontos ou nós de uma grande rede de enunciados que usamos para comparar teoria e realidade. Isso significa que uma frase teórica é levada ao teste da experiência apenas em conjunto com outras frases (sejam frases de outras teorias mais gerais, ou descrições do modo correto para o uso de instrumentos científicos, ou ainda descrições das evidências, etc.). Essa tese ficou conhecida como *holismo confirmatório* e, em virtude dela, o reducionismo mostra-se inviável: não há como determinar o conteúdo empírico de uma frase a um único relato da experiência empírica. Se a promessa do reducionismo era não apenas afastar o risco de equivaler enunciados teóricos à metafísica, mas também de explicar o que faz com que um enunciado sintético seja verdadeiro, então a crítica ao segundo dogma mostra que explicar a distinção analítico-sintético através da noção de *sinteticidade* não é menos problemática do que explicá-la partir da noção de *analiticidade*.

Isso tudo reforça o ponto que vimos na seção anterior, contra a distinção analítico/sintético, mas é a consequência epistemológica desse argumento que mais nos interessa aqui. Se o holismo confirmatório estiver correto, isto é, se *toda* a rede de enunciados que endossamos é confirmável pela experiência empírica, então *toda* a rede de enunciados que endossamos é também *refutável* pela experiência empírica. Ou seja, pode ser o caso que testar *T* com *E* leve-nos a corrigir não *T*, nem o relato das evidências *E*, mas algo que pensávamos ser incorrigível, *como o nosso suposto conhecimento a priori*. Como vimos, o conhecimento *a priori* é tipicamente concebido como um conhecimento obtido independentemente da experiência. Mas — e isso é uma consequência do que Quine teria mostrado — até mesmo o mais cristalino conhecimento supostamente *a priori*, como o de que 2+3=5, pode ser revisto pela experiência. É claro, apenas uma experiência terrivelmente cataclísmica teria força suficiente para afetar nossa confiança nos conhecimentos da lógica e da matemática, por

exemplo, mas essa possibilidade não é nem pode ser excluída de antemão. Em resumo, o argumento aqui é que, se um conhecimento é obtido *a priori*, ele independe da experiência, e se todo conhecimento que obtemos pode ser corrigido pela experiência, então não há conhecimentos que independem da experiência, isto é — ao contrário do que supunha a tradição — não há conhecimentos *a priori*. Se esse for o caso, todo o nosso suposto conhecimento está à mercê de experiências, ou seja, todo conhecimento é *a posteriori*. Note que até mesmo os empiristas clássicos da modernidade pensavam que pelo menos o conhecimento matemático é obtido *a priori* (com a exceção notável do filósofo inglês John Stuart Mill), de modo que o empirismo resultante dos *Dois Dogmas* é muito mais radical do que o empirismo moderno.

Há, no entanto, algumas estratégias de resposta a essa crítica. Uma delas consiste em argumentar que, embora o conhecimento *a priori* seja *obtido sem a experiência*, ele ainda pode ser *revistado pela experiência* (o que às vezes chamamos de *a priori banana*). Ou seja, o conhecimento *a priori* independe da experiência em um sentido (no que concerne à sua obtenção), mas depende, ou pode vir a depender, da experiência em outro sentido (no que concerne à sua revisão). Com efeito, porque deveríamos supor que o conhecimento *a priori* seria, por assim dizer, "anabolizado": absolutamente irrevogável, como supunha a tradição? A ideia do *a priori* revogável pela experiência é uma ideia interessante, mas ela só faz sentido se deslocarmos a distinção entre *a priori* e *a posteriori* do conceito de *conhecimento* para o conceito de *justificação para crer*. Porque, é claro, um conhecimento não pode ser adequadamente revisado. Se eu *sei* que 2+2=4 ou se eu *sei* que estou em Salvador ou se eu *sei* que hoje é segunda-feira e assim por diante, qualquer pretensa revisão desses conhecimentos seria um *engano*. Se eu sei de fato, e não apenas *pareço* saber, então não posso estar errado, tampouco posso revisar corretamente meus conhecimentos. Se eu devo revisar um suposto conhecimento com base em novas descobertas, é porque eu não sabia até aquele momento, não obstante a aparência de conhecimento. Por conta disso,

temos de pensar em *justificações a priori para crer*. Posso estar muito bem justificado a crer que isso-e-aquilo, posso ter excelentes evidências que até então foram incontestadas — mas basta que uma nova evidência contrária, forte o suficiente, seja incorporada às evidências das quais disponho para que eu perca minha justificação para crer que isso-e-aquilo. Ou seja, justificação para crer é revogável (ou derrogável) pela nova experiência, e somente assim podemos entender o que significa dizer que aquilo em que eu acredito *a priori* pode ser revisado pela experiência empírica.

O saldo dos argumentos elencados por Quine contra o que ele chamou de dogmas do empirismo pode ser visto por uns como uma confusão conceitual que desnecessariamente abala tudo aquilo que parecia óbvio e, de certo modo, até mesmo seguro e confortável na filosofia pregressa. Se Quine estiver correto, nós temos de lidar com incertezas à primeira vista profundamente desagradáveis: será que não há diferença substancial entre enunciados filosóficos e científicos? O que é fazer filosofia nessa nova perspectiva? Para estudar filosoficamente o conhecimento, por exemplo, eu devo me informar sobre neurociências e psicologia cognitiva? E nosso conhecimento — seria ele todo vulnerável a descobertas empíricas, mesmo aquele que habita, ou melhor, que supostamente habita o reino cristalino da lógica? O cenário resultante pode parecer desolador, de fato, mas isso também faz dele intelectualmente fértil, porque representa a possibilidade de integração da filosofia em uma perspectiva científica ampla. Com efeito, depois de *Dois Dogmas*, muitos filósofos se declararam naturalistas, e talvez o melhor argumento em favor dos resultados obtidos por Quine seja uma inegável constatação histórica: há muita filosofia empiricamente informada e empiricamente relevante sendo feita em altíssimo nível agora mesmo, enquanto nós conversamos. Os puristas podem bater o pé e insistir que isso já não é mais digno do nome de filosofia, mas se for uma questão meramente nominal, então podemos chamar de outra coisa — podemos chamar de ciência se isso os acalma. Nenhum naturalista perderia o sono por causa desse novo título para sua disciplina.

Interlúdio: como funcionam nomes?

Quine mostrou que as distinções analítico/sintético e *a priori*/*a posteriori* não são tão óbvias quanto podem parecer à primeira vista (ainda que talvez seja possível defendê-las com novos argumentos). Mas e quanto à distinção entre necessário e contingente? Aqui não vamos ver uma crítica à viabilidade dessa distinção, mas sim uma crítica à associação entre verdade necessária e conhecimento *a priori*, bem como à associação entre verdade contingente e conhecimento *a posteriori*. Embora pareça óbvio que um conhecimento obtido *a priori* possa apenas ter como conteúdo uma frase cuja verdade é necessária (ao passo que um conhecimento *a posteriori* possa apenas ser sobre verdades contingentes), essas associações serão abaladas se sucedermos em mostrar que, por um lado, há conhecimentos *a priori* de verdades contingentes e que, por outro, há conhecimentos *a posteriori* de verdades necessárias. Para isso, no entanto, precisamos entender sobre o funcionamento de nomes.

Nomes são entidades linguísticas que nos permitem, de certo modo, "apontar" para o mundo através da seleção do indivíduo nomeado. Para evitar metáforas, vou usar o termo mais geral que compreende esse gesto de "apontar" linguístico, a saber, a *designação*. Por exemplo, o nome 'Ostara' designa a gata preta, enquanto o nome 'Klaus' designa o gato branco. Ao proferir, por exemplo, o enunciado 'Ostara está sobre a mesa' eu uso a linguagem para descrever ou representar uma fatia da realidade, o estado de coisas que consiste em Ostara estar sobre a mesa. Além disso, um nome tipicamente ocorre na posição de sujeito gramatical de uma frase, por exemplo, em 'Klaus está no quarto'. Visto que Klaus, o indivíduo sobre o qual é essa frase, é designado pelo nome, e visto que o nome ocorre na posição de sujeito gramatical daquela frase, parece que o *sujeito lógico* (aquilo sobre o que é a frase) é igual ao seu *sujeito gramatical*. À primeira vista, tudo isso é incontestável, mas o assunto fica mais complicado se nos perguntarmos como nomes designam seus objetos ou indivíduos. Ou seja, como, afinal de contas, funcionam os nomes?

Gottlob Frege, o filósofo visionário que mencionamos no capítulo anterior, deu uma resposta exemplar a essa pergunta. Ele desenvolveu a influente ideia de que as expressões de uma linguagem são compostas por *sentido* e *referência* — assim como frases também o são (mas não precisamos pensar nisso agora, vamos focar na aplicação da distinção sentido/referência a nomes). Para entender o que Frege quer dizer com isso, considere a seguinte história verídica (ou, ao menos, que possui alguns detalhes verídicos): há alguns anos, eu conheci um notório professor de filosofia com uma atuação distinta em filosofia da linguagem. Chamemo-lo de João Carlos porque, bem, esse é seu nome. Naquele momento, eu formei a crença de que João Carlos é um especialista em filosofia da linguagem. Mais recentemente, quando soube que começaria a trabalhar na Universidade Federal da Bahia, eu provavelmente formei a crença de que o reitor é meu novo chefe. Vamos supor que eu ainda mantivesse a crença inicial a respeito de João Carlos. O que é interessante é que, por algum tempo, eu ainda não sabia que João Carlos era de fato o reitor dessa instituição. Assim, vamos supor que em algum momento eu tenha mantido as duas seguintes crenças:

1. João Carlos é um especialista em filosofia da linguagem.
2. O reitor é meu novo chefe.

Vamos imaginar que nesse mesmo tempo eu tivesse duvidado da seguinte afirmação, a saber:

3. João Carlos é meu novo chefe.

Mas, se João Carlos é o reitor dessa instituição, e se o reitor é meu novo chefe, então João Carlos é meu novo chefe. Eu estaria cometendo uma contradição ou violando alguma norma da racionalidade ao negar ou duvidar de 3, tendo aceitado previamente 1 e 2, e dado o fato de que João Carlos é de fato o reitor?

Parece-me que não, assim como pareceu a Frege mais há mais de um século. Mas como eu não estaria em contradição? Poderia parecer que há alguma falha da minha parte, porque se é verdade que João Carlos é o reitor, é verdade que o João Carlos é meu novo chefe! No entanto, segundo Frege, o que ocorre em um caso desse tipo é que, embora as expressões 'João Carlos' e 'reitor [da Universidade Federal da Bahia]' designem o mesmo indivíduo — ou seja, embora elas tenham a *mesma referência*, apontem para a mesma coisa no mundo — elas designam esse indivíduo de modo diferente. A escolha de uma expressão em detrimento da outra para designar João Carlos afeta o modo como o designamos, ou melhor, essas expressões *possuem sentidos diferentes*. Em uma perspectiva um pouco mais abstrata, podemos dizer que uma referência *R* pode ser designada através de sentidos diferentes *S* e *S**, de tal modo que uma pessoa que usa *S* para designar *R* não necessariamente se encontra em condições de usar *S** para o mesmo efeito. Assim, essa pessoa pode ao mesmo tempo, e sem falha de racionalidade, atribuir uma qualidade *Q* a *R* usando o sentido *S* para designá-la e negar a mesma qualidade *Q* de *R* usando o sentido *S** para designar *R*.

Frege lança luz sobre a distinção entre sentido e referência com a seguinte metáfora: imagine que você tem diante de si uma luneta que permite ver um corpo celeste qualquer, digamos, o planeta Vênus. O planeta ele mesmo, nessa metáfora, é a *referência*. O jogo de espelhos que compõe a luneta e que permite a magnificação da imagem é o *sentido*, ou o modo como a referência *é* apresentada ao observador. É claro, nessa metáfora, há alguém observando Vênus através da luneta. É possível que existam outras lunetas apontando para Vênus (diferentes sentidos designando a mesma referência). Mas a ideia principal aqui é que qualquer observador que ocupasse a mesma posição que você ocupa quando usa a luneta receberia exatamente a mesma imagem de Vênus — ou seja, o sentido ou modo de apresentação da referência é independente das peculiaridades do observador, ele é objetivo. É claro, quando usamos uma expressão qualquer podemos ter várias, potencialmente infinitas ideias subjetivamente associadas

à expressão. Por exemplo, quando eu digo que Ostara está sobre a mesa, o nome 'Ostara' está associado a várias memórias que eu possuo, mas que outra pessoa, menos familiar com a minha gata, não possuiria ao proferir as mesmas palavras que eu. Mas a metáfora da luneta mostra que isso é irrelevante para que o nome designe a referência através de um sentido específico.

A resposta de Frege à questão sobre como nomes nomeiam, portanto, passa pela distinção entre sentido e referência. Além disso, Frege parece sugerir que o sentido de um nome como 'Klaus' é uma *descrição definida,* como, por exemplo, 'o gato branco'. Uma descrição definida é uma entidade linguística caracterizada por um artigo definido e pela atribuição de uma ou mais propriedades a um indivíduo. Mas Frege não usa essas palavras e nunca explora essa noção com mais clareza. Foi Bertrand Russell que examinou cuidadosamente a noção de *descrições definidas* no começo do século passado. Filósofo, matemático, lógico, historiador e ativista político britânico, Russell é considerado, junto de Frege, Moore e Wittgenstein, um dos fundadores da filosofia analítica. Mais do que isso, Russell foi uma figura emblemática durante maior parte do século XX, e seu legado humanitário transcende suas contribuições filosóficas (que são brilhantes) — como quando foi preso em protesto à participação inglesa na primeira guerra mundial e preso também, um pouco mais tarde (dessa vez com 89 anos), em protesto contra o armamento nuclear das nações em 1961.

Antes de Russell, Frege parece ter pensado que uma descrição definida tem uma estrutura lógica simples, idêntica à estrutura de um nome. Isso é, de fato, plausível: dizer 'Klaus está no quarto' não parece diferente de dizer 'o gato branco está no quarto'. Circunscrito o contexto em que o nome e a descrição são usados, ambos enunciados são obviamente sobre o mesmo indivíduo, o gato Klaus. À primeira vista, portanto, pode parecer que o nome 'Klaus' e a descrição definida 'o gato branco' funcionam da mesma maneira. No seu artigo *Da Denotação,* de 1905, Russell mostrou que esse não era o caso, pois uma descrição definida tem uma estrutura lógica complexa que

seleciona um indivíduo no mundo (se de fato seleciona) através da satisfação de condições. Na análise de Russell, a descrição definida 'o gato branco' quer dizer que *há um único gato que é branco*, ou seja, ela seleciona um indivíduo no mundo através da satisfação das condições (1) ser um gato, (2) ser branco e (3) existir somente um gato branco. Se há mais de um gato branco nesta casa, ou se o gato não é branco, mas amarelo; ou se é uma calopsita e não um gato, então a descrição 'o gato branco' não é satisfeita por nenhum indivíduo. Ou seja, embora descrições definidas pareçam funcionar como nomes, são na verdade entidades lógicas bem diferentes, que designam indiretamente através da satisfação de condições. Já o que Russell entende por *nomes logicamente próprios* (um termo técnico) são entidades lógicas que se referem *diretamente* ao indivíduo nomeado.

Por um tempo, Russell chegou a pensar que nomes próprios da linguagem natural (como 'Klaus', 'Ostara', 'João', 'Maria', etc.) são nomes logicamente próprios, ou seja, que eles se referem diretamente, não através de descrições. Ele mudou de ideia, no entanto, quando notou que, para nomear diretamente um objeto, é preciso estar em contato imediato com esse objeto. Contudo, é sempre possível usar um nome na ausência do objeto nomeado. Pense, por exemplo, na frase 'Bismarck unificou a Alemanha'. Entendemos perfeitamente o que essa frase quer dizer, não obstante o fato de que nunca tivemos contato com o notório estadista germânico. Mas como isso seria possível? A resposta de Russell é simples: o nome próprio da linguagem natural 'Bismarck', apesar das aparências, não é um nome logicamente próprio, e sim a *abreviação* de uma descrição definida, como 'o maior estadista alemão do século XIX'. O mesmo vale para o nome 'Alemanha', que abreviaria uma descrição definida como: 'país com a maior população da União Europeia'. Entendemos o que a frase 'Bismarck unificou a Alemanha' quer dizer porque podemos *analisar* os nomes próprios da linguagem natural que nela ocorrem em descrições definidas, e repetimos esse procedimento para os outros nomes que surgirem nas novas descrições que resultarem da análise. A ideia é que assim chegaríamos, literalmente, em *última análise*, a termos

com os quais temos contato direto. Essa é a ideia fundadora da filosofia como uma *análise lógica da linguagem*, um projeto tão influente que deu origem ao nome 'filosofia analítica'. Se, portanto, é possível que um nome não se refira a entidade alguma, se é possível que aquilo que é nomeado não exista, não obstante o nosso uso significativo do nome; então um nome próprio da linguagem natural é, na visão tardia de Russell, uma descrição definida abreviada. Como é sempre possível que o objeto ao qual pretendemos nos referir através de um nome próprio não exista, pois é sempre possível que estejamos enganados (pense num cenário cético cartesiano), então os nomes próprios da linguagem natural *nunca* são nomes logicamente próprios. E quais seriam os nomes logicamente próprios, os nomes que não abreviam descrições definidas, mas que se referem diretamente em razão do usuário do nome estar em contato direto com o objeto nomeado? Russell sugere: 'isto' — seguido de um gesto de apontar — e, talvez, 'eu'.

A ideia de que nomes da linguagem natural abreviam descrições definidas é conhecida como *descritivismo* em filosofia da linguagem. Segundo o descritivismo, a relação semântica que estamos chamando até aqui de designação é, no caso dos nomes próprios da linguagem natural, uma relação de *denotação*. Um nome denota um indivíduo, nessa concepção, porque ele abrevia uma descrição definida (ou um conjunto de descrições definidas) unicamente satisfeitas pelo indivíduo em questão. Essa ideia é especialmente interessante porque resolveria algumas das nossas perplexidades filosóficas genuínas. Considere a seguinte frase: 'Humpty Dumpty está sentado no muro'. Você entende o que ela quer dizer, sobretudo se você leu *Alice no País das Maravilhas* de Lewis Carroll. Mas como seria possível entender essa frase visto que Humpty Dumpty *não existe*? Humpty Dumpty é uma invenção, uma entidade ficcional criada pelo autor do livro. Portanto, o nome 'Humpty Dumpty' é, por assim dizer, vazio. Se o nome é vazio, como pode a frase fazer sentido? A resposta descritivista é engenhosa: podemos entender perfeitamente o que significa dizer que Humpty Dumpty está sentado no muro porque o nome próprio

nesse enunciado abrevia uma descrição definida, como 'existe uma única pessoa ovoide com tendências hegelianas para o uso arbitrário da linguagem', ao passo que a frase toda, por sua vez, diz que essa pessoa está sentada no muro. É claro, é uma frase *falsa* porque Humpty Dumpty afinal de contas *não existe*, mas podemos entender perfeitamente o que ela quer dizer.

Mas será mesmo que um nome próprio da linguagem natural abrevia uma descrição definida? Ou ainda: será correto que um nome designa pela satisfação de condições, e não diretamente?

O trabalho inovador da brilhante filósofa norte-americana Ruth Barcan (mais tarde, Ruth Barcan Marcus) em 1946 foi crucial para mostrar que nomes não designam do mesmo modo que descrições definidas. Antes dela, os operadores modais, isto é, os operadores de necessidade e de contingência, não eram de grande interesse para os lógicos e filósofos da lógica. Frege e Russell, por exemplo, não se preocupavam com a pergunta sobre se algo é necessariamente ou contingentemente verdadeiro, ou se algum objeto tem uma propriedade necessária (ou essencial, no vocabulário aristotélico) ou contingente (acidental). Para aqueles filósofos, questões modais não eram importantes porque eles se interessavam por filosofia da matemática — afinal de contas, plausivelmente, toda verdade matemática é necessária (salvo, por exemplo, para empiristas radicais). Quine também era um pensador que ignorava a relevância do operador de necessidade porque esse operador estaria ligado à noção de analiticidade (que, afinal de contas, não faz sentido para Quine). Ruth Barcan mudou tudo ao argumentar, por exemplo, que, para qualquer coisa x, se x necessariamente tem a propriedade F, então necessariamente todo x tem a propriedade F. ('x' aqui é o que se chama de variável, um símbolo que não designa nada em especial).

Contra a ideia de Barcan, e na tentativa de mostrar que seria absurdo falar de propriedades necessárias — ou seja, falar de essências como os antigos as concebiam —, Quine apresenta o seguinte argumento:

1. 9 é necessariamente maior do que 7.
2. 9 é o número dos planetas.
3. Portanto, o número dos planetas é necessariamente maior do que 7.

A suposição de Quine aqui é que podemos substituir '9' por 'o número dos planetas' porque ambas expressões designam do mesmo modo. Isso é, como vimos, uma marca do descritivismo. Quine continua: é falso que o número dos planetas é necessariamente maior do que 7, pois o número dos planetas depende de como as coisas são no mundo, e é perfeitamente possível que o nosso sistema solar tivesse 7 ou menos planetas. Isso teria mostrado, segundo Quine, o absurdo em falar sobre propriedades necessárias. Agora, o interessante é que o argumento quineano pode ser interpretado de outra maneira, uma interpretação que é muito mais interessante: na primeira premissa, nós temos os nomes '9' e '7', que, segundo Barcan, referem-se como etiquetas (isto é, diretamente) aos números 9 e 7 respectivamente. Diferentemente, a expressão 'o número dos planetas' é uma descrição definida. Dizer, por exemplo '9 é o número dos planetas' não é afirmar uma *identidade* entre os referentes de '9' e de 'o número dos planetas', é atribuir uma propriedade ao número 9, isto é, de ser o número de planetas. Ou seja, substituir '9' por 'o número de planetas' falha porque não são entidades linguísticas com a mesma natureza semântica, isto é, elas designam de modos diferentes.

A ideia original de Marcus, segundo a qual nomes não podem ser substituídos por descrições quando usamos o operador de necessidade, foi desenvolvida magistralmente por Saul Kripke na série de palestras que ele ministrou em 1970 em Princeton com apenas 30 anos. Em 1980, essas palestras foram publicadas com o título *O Nomear e a Necessidade* (em inglês, *Naming and Necessity*), que seguramente é um dos livros mais importantes da filosofia, não obstante a eventual dificuldade de identificar exatamente os argumentos de Kripke em virtude de tratar-se majoritariamente das transcrições de suas falas.

O argumento principal de Kripke contra o descritivismo é bastante simples, mas ele envolve alguns exercícios imaginativos. Primeiro, vamos supor, de acordo com o descritivismo, que um nome próprio da linguagem natural abrevia uma descrição definida. Vamos supor que 'Platão' abrevia a descrição definida 'o maior filósofo da Antiguidade Clássica'. Ou seja — diria o descritivista —, quando alguém usa o nome 'Platão' em uma frase, essa pessoa está, com efeito, referindo-se a Platão através da descrição definida 'o maior filósofo da Antiguidade Clássica'. Como essa descrição é unicamente satisfeita por Platão, um enunciado como 'Platão escreveu o diálogo Teeteto' seleciona o indivíduo Platão em virtude daquela descrição. Mas as coisas poderiam ser diferentes. *É possível que* Platão não tenha sido o maior filósofo da Antiguidade Clássica. Mesmo que no nosso mundo, no mundo atual, ele de fato tenha sido o maior filósofo da Antiguidade Clássica (para o desgosto de aristotélicos), vamos imaginar um cenário em que outro pensador tivesse ocupado tão distinta posição. Poderia ter sido, digamos, um sujeito chamado Dikaiópolis. Vamos imaginar que Dikaiópolis escrevera uma obra espetacular e revolucionária, tremendamente influente na antiguidade, que eventualmente foi perdida. Assim, é possível que a descrição 'o maior filósofo da Antiguidade Clássica' denote Dikaiópolis, e não Platão. Mas, o que é importante, *nós não diríamos que o nome* 'Platão' *denotaria Dikaiópolis*! Com efeito, o enunciado 'Platão escreveu o diálogo Teeteto' continuaria sendo um enunciado *sobre Platão*, não sobre Dikaiópolis, mesmo no cenário alternativo em Platão não tivesse sido o maior filósofo da Antiguidade Clássica. Portanto, é falsa a nossa suposição inicial, segundo a qual um nome próprio da linguagem natural abrevia uma descrição definida.

Kripke explica o que acontece em um caso assim através do conceito de *designador rígido*. Um designador rígido é um termo que designa o mesmo indivíduo em todos os mundos possíveis em que o indivíduo existe. Nomes próprios da linguagem natural são designadores rígidos por excelência. Isso fica claro no exemplo acima, pois o que nos permite falar sobre Platão em um mundo possível em que

ele não é o maior filósofo da Antiguidade Clássica é precisamente o nome 'Platão'. Se Platão sequer tivesse sido um filósofo, mas, imaginemos, um halterofilista, e se a Academia de Platão fosse apenas um lugar frequentado pelos jovens atenienses para ganhar uns músculos e "ficar monstros", como se diz por aí — ainda assim é de Platão que estamos falando ao imaginarmos esse mundo possível. É claro, esse resultado depende da suposição de que mundos possíveis são estipulados a partir do mundo atual, por isso não faz sentido perguntar se Platão (filósofo) é idêntico a Platão (maromba), apesar das suas diferenças, pois nós estamos falando de quem Platão poderia ter sido, mas não é. Por outro lado, como esses exemplos mostram, descrições definidas geralmente não são designadores rígidos, porque elas denotam indivíduos a partir da satisfação de condições, o que por sua vez envolve propriedades que um individuo poderia não ter. Platão poderia não ter sido o maior filósofo da Antiguidade Clássica, como no exemplo anterior, e então a descrição em questão não o denotaria. Descrições definidas são, diz-se às vezes por aí, designadores *flácidos*.

É possível resumir o resultado de Kripke da seguinte maneira: nomes e descrições definidas não são intercambiáveis em contextos em que ocorrem os operadores de necessidade e possibilidade, ainda que atualmente um nome e uma descrição designem o mesmo indivíduo. A teoria de que nomes são designadores rígidos é conhecida como teoria da referência direta. Desse modo, ao propor uma teoria da referência direta, Kripke amplia o que Russell havia mostrado em *Da Denotação*, a saber, que descrições definidas e nomes não funcionam da mesma maneira. Nomes referem-se direta e rigidamente (selecionando o mesmo indivíduo em todos os mundos possíveis em que existe), enquanto descrições definidas denotam através da satisfação de condições. Pelas razões que nós vimos, Russell mais tarde passou a considerar que nomes próprios da linguagem natural abreviam descrições definidas, o que Kripke rejeita através do exame de como nós respondemos quando somos expostos a certos casos imaginários, que na filosofia são chamados de experimentos mentais. Ou

seja, nomes próprios da linguagem natural, para Kripke, são como o que Russell chamou de nomes logicamente próprios.

E isso é só o começo das descobertas interessantes apresentadas por Kripke.

Necessário *a posteriori* e contingente *a priori*

Se nomes próprios da linguagem natural não são como descrições definidas nem as abreviam, então afirmar 'Marco Túlio é Cícero' é diferente de afirmar 'Marco Túlio é o autor das Catilinárias'. O primeiro enunciado tem a forma de uma *identidade*, enquanto o segundo é uma *atribuição*. Com efeito, nós vimos, Marco Túlio continuaria sendo ele mesmo ainda que outra pessoa tivesse escrito as Catilinárias. Mas, visto que nomes próprios da linguagem natural são designadores rígidos, então quando se enuncia uma identidade verdadeira, como no nosso primeiro exemplo, os nomes 'Marco Túlio' e 'Cícero' necessariamente se referem ao mesmo indivíduo em todos os mundos possíveis em que esse indivíduo existe. Desse modo, uma afirmação de identidade, se verdadeira, é *necessariamente verdadeira*.

Isso, eu acredito, é incontroverso. Mas o que é interessante mesmo é que o conhecimento de que os nomes 'Marco Túlio' e 'Cícero' referem-se à mesma pessoa é uma descoberta *a posteriori*. A não ser que você seja moderadamente bem informado(a) sobre política romana, eu imagino que você provavelmente não sabia que Marco Túlio é Cícero até ler o parágrafo anterior. Talvez você soubesse que Marco Túlio foi um importante político romano, e que Cícero foi um filósofo associado aos céticos acadêmicos. Mas, e essa era minha aposta com a escolha desse exemplo, você veio a saber *a posteriori*, através da leitura deste livro, uma verdade necessária, a saber, que Marco Túlio é Cícero. Assim, a teoria da referência direta proposta por Kripke sob o plano de fundo das descobertas de Ruth Barcan contra o descritivismo mostra que verdades necessárias podem ser descobertas *a posteriori*.

E quanto ao contingente *a priori*? Bom, para entender esse ponto, considere novamente as razões pelas quais Russell concluiu que um nome próprio da linguagem natural não se refere diretamente ao indivíduo nomeado, mas apenas através de descrições abreviadas ou implícitas no conteúdo do nome. Como eu não estou em contato imediato com Sócrates, eu me referiria a ele apenas através de descrições, mesmo em enunciados como 'Sócrates foi um filósofo ateniense', em que não parece haver descrições definidas ('um filósofo ateniense' é uma descrição *indefinida* que não seleciona Sócrates especificamente). Como seria possível então, se Kripke estiver correto ao defender a teoria da referência direta, referir-se diretamente a um indivíduo com quem não é possível ter contato imediato? A resposta elaborada por Kripke envolve a ideia de uma *cadeia causal de transmissão de um nome*. Nós nos referimos a Sócrates porque existe uma cadeia ou corrente de usuários do nome, ela inicia com os contemporâneos de Sócrates e chega até nós, e isso preserva a referência original quando usamos o nome hoje em dia. Ou seja, o uso de um nome às vezes, ou até mesmo geralmente, é como se estivéssemos em "contato à distância" com o indivíduo nomeado.

Mas, é claro, há ainda um problema residual. Russell, nós vimos, rejeita que os nomes refiram-se diretamente porque sempre podemos estar enganados acerca de objetos com os quais não temos contato direto. Mesmo que agora eu tenha a impressão de estar vendo a gata Ostara sobre a mesa, *eu poderia estar enganado* sobre sua existência aqui e agora — pense novamente em um cenário cético como imaginado por Descartes. Se não existisse nada a que o nome 'Ostara' refere-se, o enunciado 'Ostara está sobre a mesa' seria sobre... *nada*! Nem sequer seria um enunciado falso, ele não teria significado. Russell não admite isso, porque parece perturbadora a ideia de que eu posso não saber sobre quem eu estou falando, ou pior, de que eu posso não saber sobre quem eu estou pensando. Mas a teoria da referência direta parece ser obrigada a admitir que uma cadeia causal de transmissão do nome pode falhar, pois talvez um elo tenha sido perdido ou talvez a cadeia tenha simplesmente surgido de

um engano, sem nomear ninguém em particular. Imagine que Sócrates nunca existiu, como às vezes se especula por aí. Imagine que ele é um personagem ficcional inventado por Platão. Se eu falasse 'Sócrates é um filósofo ateniense', então, eu estaria me referindo a ninguém, porque não haveria um indivíduo que ancorasse, por assim dizer, a cadeia causal de uso do nome. Consequentemente, na teoria da referência direta, é possível que eu não saiba sobre quem estou pensando. O preço para evitar esse problema é alto demais, a saber, negar que tenhamos contato direto com o mundo exterior, como faz Russell, porque sempre podemos estar errados a respeito do que transcende nossa experiência imediata. Perder o mundo ou perder o conhecimento sobre o conteúdo dos nossos pensamentos é uma escolha difícil.

De qualquer modo, segundo a teoria causal da referência, quando iniciamos uma cadeia causal de transferência do nome, temos um ato de *fixação da referência.* A fixação da referência é um ato de "batismo lógico", por assim dizer. Em algum momento foi dito 'este é Klaus', e então o nome pôde passar a ser veiculado para referir-se ao gato Klaus por outro falantes, alguns dos quais nunca tiveram contato imediato com Klaus. Para fixar uma referência, podemos inclusive usar uma descrição definida, o que não quer dizer que a descrição definida faz parte do conteúdo do nome. Eu poderia ter dito 'Klaus é o gato branco' no ato de fixação da referência. Note que é possível que Klaus mais tarde tivesse se tornado mais amarelado e, com isso, deixado de ser o gato branco. A propriedade de ser o gato branco, portanto, é contingente. Mas, e isso é interessante, no ato de fixar o nome de Klaus através da descrição definida 'o gato branco', eu sabia algo *a priori* sobre Klaus, a saber, que ele é branco.

Um exemplo melhor é discutido pelo próprio Kripke: no século XIX, o astrônomo francês Le Verrier previu matematicamente a existência do planeta Netuno ao constatar irregularidades na órbita de Urano. Imaginemos que ele tivesse proferido, antes de qualquer pessoa observar Netuno, as seguintes palavras: 'Netuno é o planeta que causa as perturbações na órbita de Urano'. Agora, repare que, em

um caso como esse, no ato de fixação da referência, Le Verrier sabia *a priori* que Netuno é o planeta que causa as perturbações na órbita de Urano. Mas, ao mesmo tempo, isso é perfeitamente contingente. Netuno poderia não ter esse efeito — Netuno pode inclusive futuramente deixar de ser a causa de tais perturbações — e, portanto, o enunciado em questão é apenas contingentemente verdadeiro. Casos em que a referência de um nome é fixada por meio de uma definição ou descrição definida, portanto, são casos de verdades *contingentes* conhecidas *a priori*.

Os resultados obtidos por Kripke em *Naming and Necessity* são eventos raros na história da filosofia, pois ao mesmo tempo são simples, elegantes e convincentes. O que é mais espetacular é que as possibilidades do necessário *a posteriori* e do contingente *a priori* haviam passado desapercebidas por toda a história das distinções em questão até as palestras de Kripke em 1970. Por isso não é exagero algum dizer que Kripke é um dos grandes gênios da filosofia e seguramente um dos maiores filósofos vivos.

Considerações finais

Neste capítulo, apresentamos as críticas de Quine contra o conceito de analiticidade e seu impacto epistemológico. Essas críticas solapam a possibilidade do conhecimento *a priori* e, com isso, motivam um naturalismo segundo o qual os enunciados da filosofia não são diferentes dos enunciados científicos. Além disso, nós vimos neste capítulo como contestar certas associações que estão no plano de fundo das distinções fundamentais que apresentamos no capítulo anterior, a saber, a ideia de que verdades necessárias podem ser conhecidas apenas *a priori* e que verdades contingentes podem ser conhecidas apenas *a posteriori*. Mostramos, então, as possibilidades do contingente *a priori* e do necessário *a posteriori*, e para isso tivemos de passar por um longo interlúdio a respeito de designação, nomes, descrições definidas, que são algumas das questões fulcrais da filosofia da linguagem no século XX.

Leituras recomendadas

Você encontra a reconstrução do argumento de *Dois Dogmas* em Sinclair (2013). Algumas consequências dessa argumentação para epistemologia são examinadas nos artigos já mencionados de Casullo (2015) e Teixeira (2014, 2015). Para uma discussão sobre a teoria das descrições definidas de Russell, veja Ferreira (2010) e também Coura (2014). O quinto capítulo de Russell (2008) trata da distinção entre conhecimento por contato e conhecimento por descrição, que é importante para sua concepção descritivista. A obra de Kripke (2012) foi traduzida para português e é essencial para entender a filosofia contemporânea. Para uma discussão sobre o contingente *a priori*, veja Ruffino (2013). Para uma discussão sobre o necessário *a posteriori*, veja Murcho (2015).

4

O mínimo de lógica

Não é incomum que iniciantes em filosofia vejam a lógica do mesmo modo como os estudantes de engenharia veem a disciplina de cálculo pela primeira vez, isto é, com horror e desespero, um olhar que mais tarde é substituído pela resignação. Mas isso é um exagero, até porque você já está ao menos parcialmente familiarizado com a lógica e, mesmo antes de ler esse livro, você seguramente até tentou fazer uso dela, mesmo sem saber. Com efeito, nós temos intuições — entendidas como disposições estáveis para julgar o que nos é apresentado — a respeito do que são bons e maus argumentos, e esse é o ponto de partida inicial para pensarmos a lógica, cujo objetivo principal é caracterizar bons argumentos. Pense, por exemplo, no seguinte argumento:

1. Todo gato é um felino.
2. Klaus é um gato.
3. Logo, Klaus é um felino.

O conectivo 'logo' (que poderia ser substituído por 'então', 'portanto' e assim por diante) indica a passagem de premissas para conclusão. É claro, na linguagem ordinária, isto é, de acordo com o modo como falamos normalmente, poderíamos apresentar o mesmo argumento começando pela conclusão. Nesse caso, teríamos algo assim: 'Klaus é um felino, porque Klaus é um gato e todo gato é um felino'. Aqui usamos outro conectivo, 'porque', que, nesse caso, indica as

premissas e que poderia ser substituído por outros conectivos como: 'pois', 'visto que', 'dado que' etc. Qualquer que seja o modo como apresentamos esse argumento, há uma intuição de que o argumento de 1 a 3 é um bom argumento. Agora compare com este exemplo:

4. Alguns gatos são felinos.
5. Todos felinos são domésticos.
6. Logo, todos gatos são domésticos.

Se você reparar atentamente, deve surgir uma espécie de sinal de alerta na sua consciência, pois intuitivamente, *esse argumento não é bom, não como o anterior*. E por qual razão? Se você respondeu constatando que é *falso* que todos os gatos são domésticos (porque há gato selvagens, por exemplo), por incrível que pareça, você deu a resposta *errada*! A resposta certa, a lógica nos mostra como veremos a seguir, é um pouco mais complicada. O argumento não é bom porque as premissas (4 e 5 acima) não estabelecem a verdade da conclusão exposta em 6, ou seja, o argumento não preserva a verdade das premissas para a conclusão. Isso independe de se a conclusão do argumento é *de fato* verdadeira ou falsa. Compare, por exemplo, com este novo argumento:

7. Todos os gaúchos são gremistas.
8. Nenhum gremista é colorado.
9. Logo, nenhum gaúcho é colorado.

Esse novo argumento também tem uma conclusão falsa, pois há muitos gaúchos que são colorados (inclusive o GreNal é unanimemente reconhecido como o maior clássico do futebol mundial). Mas, diferentemente do argumento anterior, esse é um *bom* argumento, segundo o critério mais básico de avaliação que apresentaremos a seguir — pois, nesse caso, a verdade é preservada das premissas à conclusão. Deve ficar claro que a lógica é uma disciplina *normativa*, a ela cabe (entre outras coisas) investigar e estabelecer os padrões da

boa argumentação. A lógica não é, portanto, uma disciplina *descritiva*, pois ela não tenta descrever os modos como de fato raciocinamos. Com efeito, é possível que exista uma divergência radical entre o modo como devemos argumentar e o modo como argumentamos na vida real, um assunto ao qual voltaremos ao fim do capítulo. Antes disso, no entanto, precisamos lançar luz sobre conceitos importantes da lógica e responder algumas questões que ficaram em aberto até aqui. Vamos começar por fazer o mínimo de justiça histórica.

Da pré-história ao passado próximo

A primeira grande descoberta da lógica na filosofia ocidental, a descoberta que talvez tenha alçado o pensamento humano a um novo patamar de clareza e rigor, foi feita por Platão, como não poderia deixar de ser. Antes de Platão, havia apenas trevas (filosoficamente). Isso pode soar um pouco exagerado, mas, se pensarmos no tipo de problema com que os filósofos anteriores estavam envoltos, veremos que não é um exagero tão grande.

Um desses filósofos era Parmênides, um pensador da ilha de Eleia, que viveu entre os séculos VI a V a.C., e cuja obra provavelmente consistia em poemas (não sabemos, afinal, porque só nos restaram fragmentos de um poema e os comentários de outros pensadores do período ou posteriores). No seu escrito mais influente, Parmênides teria chegado a uma conclusão estarrecedora: *é impossível dizer o falso!* Seu argumento parece ter sido mais ou menos o seguinte: o ser, tudo aquilo que existe, é. O não-ser, o inexistente, não é. Falar um enunciado falso seria referir-se ao não-ser, falar do que não existe, e isso é patentemente impossível, porque, ao falarmos sobre alguma coisa, referimo-nos a ela e, portanto, ela existe. Portanto, é impossível dizer o falso.

Se você achou esse argumento estranho, ou se teve dificuldade para interpretá-lo algumas vezes, saiba que não é culpa sua. Há de fato algo muito estranho no uso da linguagem aqui, algo que diz respeito ao verbo 'ser'. Platão foi o primeiro a notar o que estava errado

com o argumento clássico de Parmênides, e ele o fez com a habitual precisão no diálogo *Sofista*. Platão argumenta, não através do personagem Sócrates, como de costume, mas através do personagem apresentado como *Estrangeiro de Eleia* (a indireta!), que as frases 'o ser é' e 'o não-ser não é' estão incompletas, elas são o que chamaríamos hoje de frases malformadas. Ao proferirmos alguma coisa, como quando dizemos que 'a água é cristalina', ligamos o sujeito do enunciado ('a água') a um predicado ('cristalina') através do verbo 'ser'. O mesmo acontece, por exemplo, quando dizemos 'a água não está limpa'. Dizer, portanto, 'o ser é' e 'o não-ser não é' consiste em proferir enunciados incompletos, porque carecem de um complemento predicativo. É como se disséssemos 'a porta é', 'Klaus é', etc., enunciados que não querem dizer nada. Os enunciados no argumento de Parmênides, portanto, não fazem sentido. E como seria possível dizer o falso, falar sobre o que não existe? Platão continua, também contra o diagnóstico de Parmênides: um enunciado falso consiste minimamente na atribuição de uma propriedade a um sujeito que não possui aquela propriedade na realidade, ou na negação de uma propriedade a um sujeito que a possui na realidade. Por exemplo, se eu falo, acerca da gata Ostara, que está sobre a mesa, 'Ostara *não* está sobre a mesa', meu enunciado é falso, porque eu nego de Ostara uma propriedade (estar sobre a mesa) que ela possui presentemente. Se eu falo, também acerca dessa mesma gata, 'Ostara é uma calopsita', eu atribuí falsamente a propriedade de ser uma calopsita à gata Ostara. Nesses casos, meus enunciados são *falsos*. Se, ao contrário, eu tivesse dito que Ostara está sobre a mesa e que Ostara é uma gata, eu teria proferido dois enunciados verdadeiros.

A descoberta mais elementar da lógica (que só seria revisada séculos depois por Frege), pois, é de que a estrutura mínima para que uma frase faça sentido — e, portanto, para que possa ser verdadeira ou falsa — é de *sujeito + verbo 'ser' + predicado*. Como o verbo 'ser', de certo modo, liga o sujeito ao predicado, diz-se que se trata de uma *cópula*, de modo que essa estrutura básica é às vezes apresentada como: *sujeito–cópula–predicado* ou, simplesmente, *S é P*. Em seguida, Aristó-

teles ampliou a descoberta platônica e efetivamente criou a disciplina da lógica — não como a conhecemos hoje, mas ainda assim basicamente sem reformas substanciais até o final da modernidade. Ele a chamava de *silogística,* em que silogismos são argumentos com certas características específicas que Aristóteles determinou por razões metafísicas. A silogística é muito mais limitada do que a(s) lógica(s) contemporânea(s), e hoje em dia ela é vista por lógicos mais como uma curiosidade histórica do que qualquer coisa. Além disso, é um infortúnio que tenha sido justamente a engessada lógica aristotélica que tenha se tornado amplamente influente na Antiguidade Clássica e na Idade Média, quando, contemporaneamente a Aristóteles, os filósofos estoicos haviam desenvolvido uma lógica muito mais versátil e interessante, semelhante à de Frege.

E por que a lógica de Aristóteles é considerada tão limitada? Algumas das razões são as seguintes: em primeiro lugar, a silogística entendia que uma frase como ‘Todo homem é mortal’ não é fundamentalmente diferente de uma frase como ‘Sócrates é mortal’. Para Aristóteles, ambas expressam uma predicação: a primeira predica a mortalidade acerca de todo homem; a segunda, acerca de Sócrates, que é um indivíduo particular. Note agora que, se Sócrates é mortal, então existe um homem mortal. Analogamente, a silogística sanciona o argumento de que, se todo homem é mortal, então existe um homem mortal. Frege mostrou, no entanto, que uma frase como ‘Todo homem é mortal’ na verdade, tem uma estrutura lógica mais complexa do que uma predicação, pois se trata de uma *condicional.* Nesse caso específico, a condicional é: *para todo x, se x é um homem, então x é mortal.* Essa frase é verdadeira mesmo que não exista nenhum homem (diz-se que é vacuamente verdadeira), de modo que não devemos inferir dela que existe um homem que é mortal. Frege havia descoberto que expressões como ‘todo’, ‘algum’ e ‘nenhum’ são na verdade *quantificadores* cuja presença modifica a estrutura lógica de frases que apenas parecem predicações na linguagem natural, e que quantificadores podem ocorrer mais de uma vez em uma frase,

como em: 'todo filósofo leu algum filósofo grego'. A lógica aristotélica simplesmente é incapaz de interpretar uma frase assim.

Em segundo lugar, para Aristóteles, uma frase como 'Sócrates é o professor de Platão' tem também a mesma estrutura das frases anteriores, a saber, a atribuição de uma propriedade, ser professor de Platão, a um indivíduo, Sócrates. Compare, com a predicação: 'Sócrates é filósofo'. Dela você pode inferir que alguém é filósofo. Mas da aparente predicação 'Sócrates é professor de Platão' você pode inferir duas coisas (a) que alguém é professor de Platão, e que (b) Sócrates é professor de alguém. Ou seja, se julgarmos pelas inferências que 'Sócrates é filósofo' e 'Sócrates é professor de Platão' permitem realizar, veremos que são tipos diferentes de enunciados. Além disso, é estranho pensar que *ser professor de Platão* é uma propriedade como *ser filósofo*, e essa é, com efeito, uma limitação que faz com que a silogística tenha dificuldade para explicar como alguns argumentos são bons. Como Frege mostra mais tarde, 'é professor de' não é um predicado como em 'é filósofo', mas é uma *relação* que ocorre entre dois indivíduos, no nosso caso, Sócrates e Platão respectivamente. Essa descoberta permite a Frege fazer inúmeras formalizações que seriam inconcebíveis na silogística aristotélica e que são essenciais para dar continuidade ao projeto de reduzir a aritmética à lógica. Considere, por exemplo, a frase: 'todo rapaz ama uma menina' (com o perdão da heteronormatividade). Essa é uma frase ambígua, que pode querer dizer que *todos os rapazes simultaneamente amam uma e a mesma menina*, ou que *cada rapaz ama uma menina em particular, que não é a mesma para todos*. A lógica fregeana facilmente explica essa diferença; a lógica aristotélica, não.

Por fim, a silogística reconhece apenas quatro tipos de frases diferentes ('todo A é B', 'algum A é B', 'nenhum A é B' e 'nem todo A é B'), e todo o silogismo deve ter apenas duas premissas e uma conclusão. Além disso, há apenas três esquemas ou figuras possíveis que determinam a organização de um silogismo. Se você fizer as contas, verá que isso gera um limite de 256 argumentos possíveis, nem todos dos quais são válidos. As lógicas contemporâneas diferentemente, são li-

vres das amarras metafísicas arbitrariamente impostas por Aristóteles à lógica e, portanto, admitem um número potencialmente infinito de argumentos válidos possíveis, o que faz delas bem mais versáteis e apropriadas para quaisquer que sejam nossos fins.

Pelas razões apresentadas acima, a lógica tal como sistematizada por Frege no final do século XIX é hoje muito mais importante do que a lógica aristotélica. As observações de Frege com respeito à natureza de quantificadores e de conceitos, inclusive, permitem-nos entender com clareza o que há de errado com o argumento ontológico apresentado por Descartes. Como comentamos no segundo capítulo, Descartes argumenta (em uma das versões desse argumento) que Deus possui todas as perfeições, e que a existência necessária é uma perfeição, donde ele conclui que Deus existe necessariamente. Contudo, como Frege mostra — o que Kant já havia observado —, *existir* não é uma propriedade, é um quantificador. Dizer que algo existe, por exemplo, que existem cavalos, é dizer que o conceito de *cavalo* não é vazio, que há indivíduos que caem sob esse conceito. Vamos supor que o conceito *cavalo* seja definido como 'mamífero, herbívoro, com crina', de modo que tudo aquilo que é mamífero, herbívoro e com crina cai sob esse conceito (uma vaca, porque não possui crina, por exemplo, não cai sob esse conceito, ou seja, uma vaca não é um cavalo). Dizer, portanto, que existem cavalos, é dizer que há no mundo indivíduos que são mamíferos, herbívoros e que possuem crina. Agora, vamos imaginar que o conceito de Deus seja definido como 'ente infinito, sumamente bom, onipresente, onipotente, onisciente', algo com que Descartes plausivelmente concordaria. *Mas a existência não é uma característica de conceitos.* Dizer que Deus existe significa dizer que algo no mundo cai sob aquele conceito, que algo é infinito, sumamente bom, onipresente e assim por diante. Ou seja, do mero exame das características do conceito, não podemos inferir que há algo no mundo que satisfaz o conceito, algo que existe com as características em questão. Portanto, o erro de Descartes no seu argumento ontológico é tratar a existência como uma propriedade das coisas, quando na verdade a existência é mais adequadamente vista como

uma propriedade de conceitos, ou simplesmente um quantificador, em termos fregeanos.

Finalmente, a verdade

Você pode ter estranhado que nós usamos amplamente as expressões ‘ser verdadeiro’ e ‘ser falso’, mas que até agora não explicamos exatamente o que é a verdade (e, por conseguinte, a falsidade). Esta é nossa tarefa presente: lançar luz sobre o conceito de verdade, que ao mesmo tempo parece tão elusivo e tão óbvio. Antes de começar, precisamos fazer algumas observações gerais.

Em linguagem coloquial, dizemos às vezes que algo é verdadeiro no sentido de ser legítimo (em oposição a falsificado) ou que uma pessoa é verdadeira no sentido de ser veraz (em oposição a ser mentirosa). Esses não são os sentidos filosoficamente interessantes de ‘verdade’. O que nos interessa agora é compreender a verdade como uma característica ou propriedade semântica de frases — o que também vale para a falsidade. Com efeito, falar em ‘enunciados’ ou ‘frases’ quando lidamos com lógica não é exatamente consensual, porque o foco da lógica não são entidades temporais, isto é, que transcorrem no tempo, como atos de fala. Por conta disso, em lógica geralmente falamos de *proposições*, pois uma proposição é o conteúdo lógico de um enunciado. É possível, por exemplo, que enunciados diferentes tenham o mesmo conteúdo lógico, como em: ‘a neve é branca’ e ‘the snow is white’. Apesar de serem enunciados diferentes, por estarem escritos em línguas diferentes, ambos descrevem um mesmo estado de coisas, que é a neve ter a propriedade de ser branca, e é isso que interessa do ponto de vista da lógica. Ademais, é possível que um mesmo enunciado seja interpretado como possuindo conteúdos diferentes, como em ‘o banco quebrou’. Dependendo de como interpretamos essas palavras ou do contexto em que elas são proferidas, esse enunciado pode expressar proposições com diferentes condições de verdade. Para evitar esse tipo de coisa, claro, idealizaremos a linguagem de tal modo a deixar tudo perfeitamente claro, o que plausivel-

mente não é possível em linguagem prosaica. Neste capítulo, portanto, eu vou usar 'proposição' ao invés de 'enunciado' ou 'frase', mas, se você preferir, pode substituir mentalmente todas as ocorrências desses termos um pelos outros, desde que você tenha em mente que há uma diferença entre um ato de fala declarativo, como um enunciado ou uma frase, e o conteúdo que esse ato expressa, uma proposição.

Ademais, na lógica clássica (da qual nos ocuparemos aqui), quando uma proposição qualquer (que podemos chamar de P por simplicidade) é verdadeira, ao negá-la, geramos uma nova proposição que diz que não é o caso que P, ou simplesmente 'não-P'. Porque P é verdadeira, a negação de P é uma proposição *falsa*. A negação desempenha essa função interessante de inverter o valor de verdade (isto é, a verdade ou a falsidade) de uma proposição. Por exemplo, dada a proposição falsa de que Klaus é um gato preto, ao negá-la, obtemos 'Klaus não é um gato preto'. Essa nova proposição é verdadeira, porque consiste na negação de uma proposição falsa. Isso é o básico, o consensual com respeito à relação entre verdade e falsidade. É claro, há casos em que *não sabemos* se uma proposição é verdadeira (ou falsa), mas isso não quer dizer que ela não seja verdadeira (ou falsa). Por exemplo, a proposição 'existe vida inteligente em outros planetas' é ou verdadeira ou falsa, ainda que não saibamos, atualmente, se ela é verdadeira ou se ela é falsa. De modo semelhante, vamos assumir por simplicidade que casos de vagueza dizem respeito a nós, não ao mundo. Alguns conceitos, como *ser calvo*, são vagos, porque não há um limite claro e bem definido entre ser calvo e ser capilarmente bem afortunado. Portanto, é possível que uma proposição como 'Fulano é calvo' seja vaga e, portanto, que habite um limiar entre a verdade e a falsidade. Para manter as coisas moderadamente sob controle, no entanto, eu vou supor que essa dificuldade é *nossa* — pois nós não temos clareza com respeito que queremos dizer com o predicado *ser calvo* — ou seja, que a vagueza não é uma característica da realidade ela mesma. Esse artifício nos permite manter que verdade e falsidade opõem-se sempre nitidamente, sem zonas cinzentas, ainda que nós às vezes sejamos incapazes de discriminá-las (quem defende que

conceitos fazem parte da realidade obviamente não aceitaria essa estratégia).

Considere agora a proposição (*P*): 'Ostara está sobre a mesa'. Há uma forte disposição para aceitar que *P* é verdadeira se, e somente se, ela *corresponde* a certo estado de coisas na realidade, a saber, o fato de que Ostara está sobre a mesa. Do mesmo modo, se na realidade Ostara não está sobre a mesa — isto é, se não há um fato ao qual à proposição *P* corresponderia —, então *P* é falsa. Essa plausível concepção sobre a natureza da verdade recebe, por razões óbvias, o nome de *correspondentismo*, e pela maior parte da história da filosofia foi a principal teoria da verdade. De acordo com essa teoria, a verdade é uma relação de correspondência entre linguagem e realidade — mais precisamente, entre uma proposição e um fato. Nessa perspectiva, uma proposição qualquer P é verdadeira se corresponde à realidade (a um fato), e falsa se não corresponde (não há fato que corresponda a *P*).

Apesar de parecer incontroverso, o correspondentismo tem alguns graves problemas. O primeiro deles pode ser constatado se considerarmos a seguinte proposição 'não existe o maior número primo'. Essa proposição, que é plausivelmente verdadeira, afirma a *inexistência* de uma entidade, e, se o correspondentismo estiver correto, ela é verdadeira porque corresponde a um fato. O problema é que esse tipo de fato parece intuitivamente diferente do fato, por exemplo, de Ostara estar sobre a mesa. A *inexistência* de uma entidade como o maior número primo sugere que se trata não de um fato normal, um estado de coisas organizado de determinada maneira, mas de um não-fato, um *fato negativo*. É claro, um correspondentista poderia evitar esse problema fazendo uma concessão substancial e limitando sua teoria a verdades empíricas (isto é, excluindo verdades matemáticas, da lógica e talvez da filosofia). Mas mesmo uma verdade empírica expressa na forma negativa parece comprometer o correspondentismo com a ideia ontologicamente desconfortável de fatos negativos, como em 'Klaus não é um gato preto'. Essa proposição é verdadeira, mas a que exatamente ela corresponde na realidade? Supostamente,

ao fato de que Klaus não é preto — ou seja, a um fato negativo. É estranho pensar que a realidade é composta de dois tipos de fatos, os "fatos positivos", como as coisas são, e os "fatos negativos", como as coisas são sem serem.

Mesmo que você não se incomode com a inflação metafísica decorrente da postulação de fatos negativos, considere outro problema, este ainda mais agudo. Pense no que significa dizer que uma coisa corresponde a outra. Pense, por exemplo, que você tem uma chave que é capaz de abrir uma fechadura específica — ou seja, que há uma correspondência entre os dentes e os sulcos da chave e o mecanismo interno da fechadura. Se não há tal correspondência entre a fechadura e a chave, é claro, essa chave não funciona. Não basta que chave e fechadura sejam semelhantes, deve haver a relação mais forte de correspondência para que a chave funcione. Entendemos o que quer dizer, nesse exemplo, que uma coisa corresponde à outra. Mas essa comparação envolve duas entidades com características similares, a chave e a fechadura. Ambas são entidades físicas, espaço-temporais, e é isso que permite que as características de uma correspondam às da outra. O problema para o correspondentismo é que estabelecer uma correspondência entre as características de uma proposição, que é uma entidade lógico-linguística, e a realidade, que é espaço-temporal, não é tão simples. A proposição 'Ostara está sobre a mesa' é composta de sujeito lógico, o nome que se refere a Ostara, a relação de *estar sobre* e o complemento, a descrição definida que denota uma mesa específica. Mas a fatia da realidade que consiste em Ostara estar sobre a mesa é composta de uma gata estar posicionada sobre um objeto e, em última análise, sobre certa organização de átomos. Visto que são coisas com naturezas distintas, pois se tratam, por um lado, de entidades linguísticas e, por outro, de corpos espaço-temporais, estabelecer uma *correspondência* entre ambas parece impossível, não obstante a plausibilidade inicial do correspondentismo. Até mesmo a ideia mais fraca de semelhança está ameaçada, pois como seria possível comparar duas entidades de naturezas *completamente* diferentes?

Esse tipo de dificuldade levou alguns filósofos, tipicamente proponentes de alguma forma de idealismo, a oferecer uma teoria da verdade alternativa ao correspondentismo. Nessa perspectiva alternativa, a verdade não é entendida como uma relação de correspondência entre linguagem e realidade, mas uma relação de *coerência* entre proposições, daí a ideia de uma teoria *coerentista* da verdade. De acordo com o coerentismo, a verdade é considerada primariamente uma característica ou propriedade de um conjunto coerente de proposições e secundariamente uma propriedade das proposições contidas nesse conjunto. Por exemplo, a verdade da proposição de que Ostara está sobre a mesa seria explicada pelo fato de que ela é coerente com outras tantas proposições, como 'a mesa é um objeto de superfície plana', 'Ostara é um gato', 'gatos tipicamente ficam sobre objetos de superfície plana', 'eu estou vendo uma gata sobre a mesa' etc. O conjunto formado por essas proposições é coerente porque todas as proposições em questão podem ser verdadeiras simultaneamente. Por outro lado, a proposição 'Ostara está voando pela sala' pode ser considerada falsa justamente porque não forma um conjunto coerente com as outras proposições.

O coerentismo é uma ideia interessante, pois tem a virtude de evitar a dificuldade, que parece fatal ao correspondentismo, de que não é possível estabelecer uma correspondência entre entidades de naturezas distintas. No coeretismo, a relação de coerência dá-se entre entidades com a mesma natureza, a saber, proposições. Por outro lado, o coerentismo enfrenta um problema específico muito pior: é perfeitamente possível que um conjunto coerente de proposições contenha somente proposições falsas. Pense, por exemplo, naquelas pessoas que se informam quase exclusivamente por mídias alternativas (mídias que não passam pelo crivo jornalístico e pela responsabilização pública que são característicos da mídia tradicional). Pegue um subconjunto qualquer das crenças de uma dessas pessoas e você verá que esse subconjunto contém certas especulações, como as proposições: 'a Terra é plana', 'a pandemia de 2020 foi uma farsa', 'uma estrela de Hollywood colocou fogo na floresta', 'não existem mudanças

climáticas', etc. Todas essas proposições *podem* ser simultaneamente verdadeiras, mas é óbvio que isso não faz delas *de fato* verdadeiras. A coerência não é, portanto, suficiente para a verdade. Como disse o ensaísta norte-americano Ralph Waldo Emerson no século XIX, *uma consistência tola é o trasgo das mentes pequenas*.

Seria possível oferecer outra teoria da verdade, uma que, ao mesmo tempo, evitasse a inflação metafísica do correspondentismo e garantisse que a verdade é uma relação entre linguagem e mundo, não entre linguagem e linguagem? Se possível, seria uma teoria que captura as virtudes do correspondentismo e evita seus problemas sem também se comprometer com a consequência contenciosa do coerentismo, a saber, de que a verdade seria independente de como as coisas são na realidade.

Uma tentativa nesse sentido é a *teoria deflacionária da verdade,* que foi esboçada inicialmente por Frege. Ao perguntar-se o que significa dizer, de uma proposição *P*, '*P* é verdadeira', Frege constatou que o termo 'é verdadeira' não acrescenta nada à asserção de que *P*. Para pegar um exemplo concreto, asserir 'é verdade que Ostara está sobre a mesa' equivale a asserir 'Ostara está sobre a mesa' (de modo semelhante, dizer 'é falso que Ostara está sobre a mesa' é equivalente a dizer 'Ostara não está sobre a mesa'). Nós não acrescentamos nada quando introduzimos um predicado 'é verdade' a uma frase a que asserimos. Nessa concepção, esvazia-se o significado de 'verdade' para além da afirmação de que as coisas são assim-e-assado, por essa razão essa teoria é dita *deflacionária*. Ou seja, a teoria deflacionária não atribui um significado ao termo 'verdade' e, portanto, evita compromissos metafísicos disputáveis sobre a natureza da verdade, diferentemente do correspondentismo, e também evita tratar a verdade como uma relação entre proposições, diferentemente do coerentismo. Mas por que, então, temos na nossa linguagem o termo 'verdade', uma expressão que não tem significado algum?

Em meados do século XX, o lógico e matemático polonês Alfred Tarski — mais um fugitivo do fascismo na Europa que encontrou abrigo no então mundo livre dos Estados Unidos — apresentou a *Con-*

venção T. A Convenção T afirma aproximadamente o seguinte: *'P' é verdadeiro se e somente se P*. Nessa convenção, você pode notar que as aspas simples no lado esquerdo do bicondicional (caracterizado pelo conectivo 'se e somente se') desaparecem, por isso se trata de um dispositivo de remoção de aspas. Considere este exemplo: a proposição 'Klaus é um gato branco' é verdadeira se e somente se Klaus é um gato branco. Segundo a teoria deflacionária da verdade, *isso é tudo que precisamos para entender o conceito de verdade*, nada mais.

O que faz com que um argumento seja bom?

Nós estamos tratando de argumentos desde o começo deste capítulo — mas, na verdade, desde muito antes de você abrir este livro, a argumentação faz parte da sua vida. Se você já tentou persuadir racionalmente uma pessoa sobre uma opinião, você já argumentou e, portanto, não é estranho a argumentos. Se você já mudou de opinião ao ser apresentado a um novo argumento, você também não é estranho a argumentos. Mas ainda não definimos cuidadosamente o que afinal de contas é um argumento.

Posto de maneira muito simples, um argumento é *um conjunto de proposições ligadas por (pelo menos) um conectivo de consequência lógica, de modo que algumas proposições são apresentadas como a consequência lógica de outras*. Temos algumas observações a fazer sobre essa definição. Falamos em consequência lógica porque há pelo menos outro tipo de consequência, a consequência causal, que não nos interessa aqui. Eu posso dizer, por exemplo, 'a gravidade causou a queda do copo'. Essa frase expressa uma noção de consequência, mas não de consequência lógica. Além disso, um argumento precisa ter pelo menos duas proposições (nesse caso, uma premissa e uma conclusão), mas não há limite para o número de proposições que pode ter, seja como premissas, seja como conclusão. Ademais, a definição acima serve para argumentos que *pretendem* que a(s) conclusão(ões) siga(m)-se da(s) premissa(s). Chamamos esse tipo de argumento de *dedutivo* — e, como veremos no capítulo seguinte, obviamente há ou-

tros tipos de argumentos que não se pretendem dedutivos, e para eles temos outros tipos de avaliação.

Agora, nós vimos na abertura do capítulo que podemos começar a pensar a lógica a partir da nossa intuição sobre o que faz com que um argumento seja bom e que isso tem a ver com a preservação da verdade. Relembre o argumento de 1 a 3 que vimos no começo:

1. Todo gato é um felino.
2. Klaus é um gato.
3. Logo, Klaus é um felino.

E agora tenha em mente este novo argumento:

10. Todo ser humano é mortal.
11. Sócrates é um ser humano.
12. Logo, Sócrates é mortal.

Ambos parecem argumentos corretos, diz-nos nossa intuição (que ainda não foi submetida a uma avaliação cuidadosa). Com efeito — e esse foi o grande acréscimo de Aristóteles à disciplina — a explicação pela qualidade positiva de ambos argumentos reside no fato de que eles têm a mesma *forma*. Se nós abstraíamos as expressões que fazem referência a pessoas e a conceitos, como 'Klaus', 'gato', 'ser humano', etc., substituindo essas expressões por letras que não se referem a nada, como '*A*', '*B*' e '*C*' abaixo, mantendo constante todo o resto, veremos que ambos os argumentos tem a mesma estrutura lógica subjacente:

13. Todo *A* é *B*.
14. *C* é *A*.
15. Logo, *C* é *B*.

A formalização que fizemos do argumento de 13 a 15 é importante porque nos permite ver claramente a sua estrutura, que é crucial

para avaliar a qualidade de um argumento de acordo com o critério mais básico da lógica, a chamada *validade*. Validade é um termo técnico que é definido da seguinte maneira: um argumento é dedutivamente válido somente se, necessariamente, se as premissas forem verdadeiras, a conclusão é verdadeira. Entendido isso, podemos ver que um argumento é dedutivamente inválido somente *se é possível que as premissas sejam verdadeiras e a conclusão seja falsa*. O argumento 13–15 nos permite avaliar claramente a sua validade, porque, se assumirmos que 13 e 14 são proposições verdadeiras (as premissas do argumento), então somos compelidos racionalmente a aceitar a verdade da proposição em 15 (seja como for que interpretemos os termos 'A', 'B' e 'C'). Compare agora com o argumento a seguir:

16. Alguns solteiros são ricos.
17. Todos os ricos são felizes.
18. Logo, todos os solteiros são felizes.

Podemos notar que esse argumento não exibe a mesma virtude que o anterior, porque é possível que as premissas sejam verdadeiras sem que a conclusão também o seja. Isso fica especialmente claro se repetirmos o procedimento de formalização que empregamos anteriormente:

19. Alguns *S* são *R*.
20. Todos *R* são *B*.
21. Logo, todos *S* são *B*.

Você pode escolher as letras que quiser aqui, conquanto mantenha as letras que escolheu constantes em todas as ocorrências dos termos originais. Fica claro, nesse exemplo, que a conclusão em 21 não se segue das premissas e, portanto, que é possível que as premissas sejam verdadeiras e a conclusão seja falsa. Ou seja, trata-se de um argumento dedutivamente inválido. Do mesmo modo, o argumento de 4 a 6 também é inválido. Relembre:

4. Alguns gatos são felinos.
5. Todos felinos são domésticos.
6. Logo, todos gatos são domésticos.

Nesse caso, assim como no argumento anterior, é possível que as premissas sejam verdadeiras e a conclusão seja falsa. Agora, é especialmente importante notar que, de acordo com a definição de validade, um argumento pode ter *premissas falsas* e ainda assim ser válido. Considere:

22. Ostara é uma gata branca.
23. Klaus é um gato preto.
24. Ostara e Klaus estão na sala.
25. Logo, há pelo menos dois gatos na sala.

Se você leu atentamente este livro até aqui, já deve saber que Ostara é uma gata preta e que Klaus é um gato branco. Portanto, as premissas 22 e 23 do argumento acima são falsas. Ainda assim, ele é tal que, se as premissas forem verdadeiras, então é necessário que a conclusão seja verdadeira. De fato, algumas premissas nesse argumento não *são* verdadeiras, mas, *se forem* (note o modo verbal), então a conclusão é verdadeira, necessariamente. Pode até mesmo ser o caso que todas as premissas e a conclusão sejam falsas (imagine que os dois gatos estão no quarto). Mesmo assim, o argumento 22–25 seria válido. Continuaria válido, inclusive, se acrescentássemos outras premissas falsas. Considere agora a seguinte questão: pode um argumento ter premissas verdadeiras e conclusão verdadeira e ser *inválido*? Aparentemente, não. Será mesmo? Atente o seguinte argumento:

26. Você está lendo este livro.
27. Logo, a Lua não é feita de queijo.

Você deve ter notado que esse argumento é inválido — o que não tem nada a ver com o fato de ele ter apenas uma premissa (é um vício

aristotélico pensar que todo argumento deve, obrigatoriamente, ter duas premissas e uma conclusão). O que faz com que ele seja inválido é que é perfeitamente possível que a premissa em 26 seja verdadeira e a conclusão em 27 seja falsa, não obstante o fato de que a conclusão é atualmente verdadeira. O que falta, em um caso desses, é uma relação apropriada de consequência lógica entre a premissa e a conclusão. Agora, preste atenção neste próximo argumento e, antes de ler o parágrafo seguinte, avalie cuidadosamente se você o considera válido:

28. Está chovendo.
29. Não está chovendo.
30. Logo, a Lua é feita de queijo.

Lembre, estamos lidando com argumentos, não com inferências que transcorrem no tempo. Assumindo que argumentos são entidades atemporais, as premissas em 28 e 29 não podem ser ambas verdadeiras. Ou seja, esse argumento é tal que não pode ter premissas verdadeiras. Além disso, a conclusão é falsa. Seria ele inválido? Surpreendentemente... *Não*! Para entender, relembre a noção de invalidade: um argumento é inválido somente se é possível que as premissas sejam verdadeiras e a conclusão seja falsa. O argumento de 28 a 30, contudo, é tal que, por conter uma contradição entre as premissas, *é impossível* que as suas premissas sejam verdadeiras. Ele pode ter uma premissa verdadeira, mas não as duas — portanto, é impossível que tenha premissas verdadeiras e uma conclusão falsa. Desse modo, apesar de parecer um péssimo argumento, ele ainda é melhor do que os argumentos inválidos que vimos anteriormente, porque 28 a 30 trata-se de um argumento válido. Também por essa razão devemos sempre ser cuidadosos em uma argumentação para não cometer uma contradição, porque, de uma contradição, tudo se segue. É o tipo de argumento válido mais barato possível.

Validade, eu venho insistindo, é o *mais básico* critério lógico para avaliar a qualidade de um argumento dedutivo, isto é, um argumento

que é estruturado de tal maneira que a conclusão pretende seguir-se logicamente das premissas. Se um argumento dedutivo não atende esse critério, ele é um mau argumento. Eu imagino que, até o momento, você ficou insatisfeito com argumentos válidos que contêm proposições falsas. Isso ocorreu provavelmente porque você está pensando no segundo critério que usamos para avaliar a qualidade de um argumento, que é chamado de *solidez* (ou *correção*). Um argumento é sólido somente se ele é *válido e possui premissas verdadeiras*. Como um exercício, revise os argumentos que apresentamos anteriormente e considere quais são sólidos. Você notará que a formalização não é suficiente para realizar esse exercício, porque a verdade, como vimos, é uma noção semântica, e a formalização consiste precisamente em abstrair o caráter semântico das proposições. É apenas porque você tem as informações de que Ostara é uma gata preta e de que Klaus é um gato branco, por exemplo, que você pode julgar o argumento 22 a 25 como *não-sólido* apesar de ser válido. Agora, como comentamos ao ver aquele exemplo, um argumento pode ser válido e ter premissas e conclusão falsas. Outro exemplo seria o seguinte:

31. Tudo que é maior do que o Sol é feito de queijo.
32. A Lua é maior do que o Sol.
33. Portanto, a Lua é feita de queijo.

Esse argumento é válido, mas não é sólido. Mas possível que um argumento *seja sólido e tenha conclusão falsa*? Pense por um momento. Se você ainda não chegou a uma resposta, relembre as definições: um argumento é válido somente se, se as premissas forem verdadeiras, a conclusão necessariamente é verdadeira. E ele é sólido somente se, além de válido, as premissas são de fato verdadeiras. Dessas definições se segue que, se um argumento é sólido, ele não pode ter conclusão falsa.

Agora, com todas essas definições em jogo, você pode ter se perguntado: por que eu precisaria saber dessas coisas para ler, entender ou até mesmo fazer filosofia? Ocorre que argumentos filosóficos tipi-

camente são dedutivos — e, obviamente, os filósofos e as filósofas que deles fazem uso pretendem que sejam argumentos dedutivamente válidos e preferivelmente sólidos também. É, portanto, crucial para examinar criticamente uma tese filosófica avaliar se as razões que a fundamentam são adequadas do ponto de vista lógico. Considere o seguinte argumento cético:

34. Se eu sei que aqui está uma mão, eu sei que não estou sonhando.
35. Eu não sei que não estou sonhando.
36. Portanto, eu não sei que aqui está uma mão.

Esse argumento é intuitivamente válido — mas é claro que nós não estamos dispostos a aceitar a conclusão em 36, não porque ela fala do conhecimento de que aqui está uma mão em particular, mas porque ameaça todo o nosso conhecimento empírico, como vimos na discussão sobre Descartes. Se quisermos, portanto, rejeitar a conclusão sem abrir mão da intuição de que o argumento é válido, devemos apresentar boas razões para rejeitar pelo menos uma das premissas. Se tivermos êxito nessa avaliação, teremos mostrado que pelo menos uma das premissas é falsa, de modo que o argumento pode ser válido, mas não é sólido. Para avançar um exame crítico de uma tese filosófica, portanto, um mínimo de lógica é necessário: reconhecer o padrão de um argumento, identificar as premissas, a conclusão, avaliar a relação lógica entre elas, etc.

Você está raciocinando corretamente?

Um dos argumentos de forma mais comum, que inclusive já apareceu neste capítulo e alhures neste livro, recebe o nome de batismo de *modus ponendo ponens* (em latim: *o modo que afirmando afirma*), mas se você o conhece deve ser pelo apelido de *modus ponens*. Considere-o a seguir, em que '*P*' e '*Q*' abreviam duas proposições quaisquer (você pode substituir essas letras por proposições específicas se as-

sim desejar, mas mantenha a substituição constante para cada ocorrência de '*P*' e '*Q*'):

37. Se *P*, então *Q*.
38. *P*.
39. Logo, *Q*.

Eu espero que não haja dúvida quanto à validade desse argumento. Considere agora que podemos fazer o caminho contrário, o chamado *modus tollens* (cujo nome completo é '*modus tollendo tollens*'):

40. Se *P*, então *Q*.
41. Não é o caso que *Q*.
42. Logo, não é o caso que *P*.

A premissa 41 pode ser mais simplesmente escrita como 'não-Q' e, do mesmo modo, a conclusão em 42 seria 'Logo, não-*P*'. Assim como 37–39, o argumento de 40–42 é válido, ainda que possa não ser tão óbvio à primeira vista. Tenha em mente agora os seguintes argumentos:

43. Se *P*, então *Q*.
44. *Q*.
45. Logo, *P*.

46. Se *P*, então *Q*.
47. não-*P*.
48. Logo, não-*Q*.

Seriam os argumentos 43–45 e 46–48 válidos? Importantemente, as premissas em 43 e 46, assim como em 37 e 40, não estão dizendo que *somente se P, então Q*. A afirmação em questão é bem mais fraca, ela diz que *se P é verdadeira, Q também o é*. Mas ela é mais fraca porque não diz que, *se Q é verdadeira, P também o é*. Ou seja, pode ser o caso

que *Q* é verdadeira sem que *P* também o seja. Desse modo, fica claro que o argumento de 43–45 é inválido. Semelhantemente, o argumento 46–48 também é inválido, porque a falsidade de *P* (como afirma 47) não diz nada a respeito da falsidade de *Q*, como se deseja concluir em 48. Cometer esses erros argumentativos, como nos últimos dois argumentos acima, consiste em raciocinar *falaciosamente* — como o nome sugere, trata-se de cometer uma *falha* lógica. Essas falácias em especial são chamadas, respectivamente, de *falácia da afirmação do consequente* e *falácia da negação do antecedente*, e são classificadas como falácias formais porque tentam, sem sucesso, reproduzir a forma de argumentos válidos, o *modus ponens* e o *modus tollens*.

Uma falácia pode ser um erro genuíno e sincero provocado pela desatenção, como pode ter sido o seu caso ao julgar os últimos dois argumentos acima como válidos. Contudo, nós frequentemente nos referimos a falácias como o mau uso *intencional* de um argumento. Como quando uma pessoa apresenta um argumento de caráter logicamente duvidoso como se fosse um bom argumento com a finalidade de persuadir (ou desautorizar) um interlocutor. Basicamente todo manual introdutório de lógica contém uma categoria extensiva de falácias, de modo que não nos é necessário passar em revista todas as falácias conhecidas (como de costume, você pode consultar as leituras recomendadas no fim do capítulo). Aqui vamos ver apenas alguns casos comuns e familiares.

Uma falácia informal muito recorrente é a *falsa dicotomia*, também chamada de *falso dilema*. Um exemplo desse tipo de falácia é o seguinte: 'ou você apoia o nosso líder, ou você é contra a pátria'. É claro, nesse tipo de fala, subentende-se que é moralmente condenável ser contra a pátria, de modo que você teria o dever de apoiar o líder. Note, contudo, que você pode discordar politicamente de alguém sem que, por conta disso, você manifeste traços antipatrióticos ou esteja associado a outro grupo político (supostamente) diametralmente oposto. Trata-se, portanto, de uma falsa dicotomia, pois você não é obrigado a ser uma coisa ou outra. Seria estúpido e perverso pensar que existem apenas duas posições políticas possíveis, visto que questões polí-

ticas são complexas e o que espectro político é nuançado. Outro caso emblemático desse tipo de falácia é o seguinte: 'ou salvamos vidas, ou salvamos a economia'. O argumento prossegue com a observação de que perderemos vidas de qualquer maneira, então seria preferível ao menos salvarmos a economia. Ora, isso é uma falsa dicotomia, porque a economia sofrerá mais gravemente com uma perda massiva de vidas. Pense, por exemplo, como a maioria dos países civilizados lidaram com a crise do COVID-19, isto é, desenvolvendo estratégias que consistem em *salvar vidas e salvar a economia*. Com isso foram propostos fundos emergenciais dedicados a pessoas carentes para que possam continuar consumindo bens básicos, além da suspensão de cobranças de aluguéis e de contas de luz e água, por exemplo.

A falácia do *argumento de autoridade* também é usada à exaustão, e você certamente já ouviu uma ou outra dessas por aí. A ideia é que uma alegação ganharia plausibilidade porque recebe o endosso de uma autoridade cuja especialidade (se possuir alguma) não é relacionada com o conteúdo da alegação. Por exemplo: 'uma marca de refrigerantes adoça seus produtos com fetos abortados, porque o professor disse isso'. Vamos supor que o "professor" em questão seja um senhor sem nenhuma formação em química industrial, ou melhor, sem nenhuma formação. Nesse exemplo, essa pessoa claramente não é uma autoridade na produção de refrigerantes, mas isso não impede que sua palavra seja invocada por um incauto em favor daquela alegação absurda. O caso teria sido completamente diferente se uma autoridade na área tivesse chegado à bizarra conclusão (o que, de qualquer modo, nunca aconteceu). Compare com outra situação muito diferente: imagine que climatologistas afirmam que há boas evidências para concluir que as mudanças climáticas são causadas pela ação humana. Nesse novo caso, estamos autorizados a acreditar nesse resultado, porque se trata do consenso entre especialistas. Note, sobretudo, que nossa autorização para crer que as mudanças climáticas são causadas pela ação humana não é prova da verdade desse enunciado. Alternativamente, podemos dizer que nós deferimos aos especialistas as boas razões para concluir que aquele enun-

ciado seja verdadeiro, mas nossa deferência não é uma razão em favor da verdade daquele enunciado.

Outra boa contendedora ao título de falácia mais popular é o *ataque à pessoa*. O conceito é simples: pretende-se desqualificar uma alegação em virtude de algo que a pessoa que a proferiu é ou fez. Por exemplo: 'os especialistas disseram que não está comprovada a eficácia de um medicamento no tratamento do novo coronavírus, mas os especialistas votaram no partido rival, portanto, é um medicamente eficaz'. Note que, para avaliar a verdade da alegação dos especialistas, é irrelevante saber em quem eles votaram ou deixaram de votar. Aprofundando um pouco nesse exemplo, se os especialistas tivessem votado no partido da pessoa que comete a falácia, continuaria sendo irrelevante para a verdade da proposição acerca da eficácia do suposto medicamento.

O ataque à pessoa é muito parecido com outra falácia muito comum, conhecida como *tu também*. Nesse caso, a pessoa que profere a falácia pretende eximir-se de culpa acerca de algo que fez ao acusar o seu interlocutor de ter feito a mesmo coisa. Por exemplo, ao ser acusado de corrupção, uma pessoa pode dizer 'você também é corrupto, então quem é você para me denunciar?'. A pretensa relação lógica aqui é obviamente falha, porque o fato de o acusador ter cometido o mesmo ato que o acusado não é relevante para retirar-lhe a culpa. Pode ser o caso que o acusador cometa a falha moral da hipocrisia — o que é certamente condenável do ponto de vista estritamente ético — mas isso por si só é insuficiente para que a acusação seja desconsiderada, tratada como falsa ou como sem sentido.

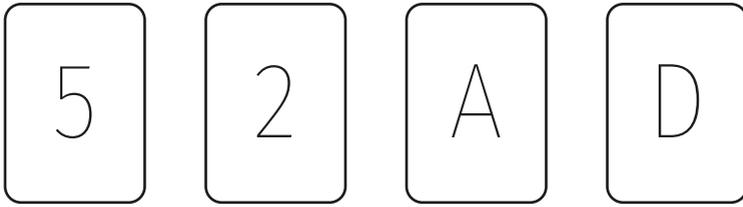
Além disso, a falácia *tu também* pode ser invocada em nome de terceiros, o que recebe o nome de *doisladismo*. Há dois tipos de doisladismo, o *doisladismo narrativo* e o *doisladismo de isenção*. O doisladismo narrativo refere-se à exigência de que os dois lados de uma história ou narrativa sejam ouvidos, o que pressupõe que toda história tenha dois lados. Ouvir as duas partes, por exemplo, de uma disputa judicial é razoável porque devemos ponderar os argumentos dos envolvidos — mas essa exigência não se traduz para *todos* os casos. Por

exemplo, recentemente um jovem comunicador foi a um canal de televisão questionar por que as autoridades davam apenas ouvidos a epidemiologistas quando o assunto é a pandemia em curso. Mas *ao que mais eles deveriam dar ouvidos? Ao vírus*? Essa seria, é claro, uma exigência irracional, fruto do doisladismo narrativo. Quanto ao doisladismo de isenção, considere este exemplo também real: em 2020, uma aglomeração de pessoas em defesa da suspensão dos nossos direitos constitucionais (por surreal que pareça!) resultou na agressão de uma transeunte que nada tinha a ver com a manifestação. Algumas pessoas defenderam os agressores da seguinte maneira: 'outros grupos políticos também já agrediram pessoas' — o que elas gostariam de continuar a dizer, mas que de fato não dizem (talvez por falta de coragem), é: 'então não há nada de errado com esta agressão em particular'. No entanto, seu enunciado é, na melhor das hipóteses, absolutamente irrelevante. Pois, mesmo que seja verdade que outros grupos políticos também já agrediram pessoas em outras ocasiões (o que é obviamente condenável), a agressão presente não se torna moralmente justificada em razão disso. Se não for verdade que outros grupos políticos cometeram agressões, por outro lado, a pessoa que enuncia isso viola uma plausível norma conversacional de que não devemos alegar falsidades. Com efeito, o doisladismo de isenção é um apelo discursivo muito comum, e frequentemente ocorre quando um dos lados de uma disputa é acusado de cometer alguma falha e deseja desviar a atenção da audiência para eventuais falhas dos seus adversários.

Estratégias que visam a desinformação e o desvio da atenção da audiência não são tipicamente classificadas como falácias. De certo modo, podemos aproximá-las das falácias tradicionais, pois tipicamente essas estratégias exibem a falta de uma relação lógica apropriada entre os enunciados proferidos ou subentendidos por agentes discursivos. A diferença é que, diferentemente de falácias, recursos desse tipo têm como objetivo principal incumbir o interlocutor de novos encargos conversacionais, pois falar asneiras e estupidezes infundadas é fácil e requer apenas criatividade (ou uma boa dose de

descolamento da realidade), ao passo que refutá-las envolve atenção aos fatos, o que toma tempo e exige rigor. Vamos supor que seu interlocutor negou que existam queimadas numa floresta tropical porque ela é úmida. Vamos supor também que as queimadas são amplamente documentadas por instituições bem reputadas. Enquanto você se esforça, por exemplo, para mostrar imagens de satélite e prepara-se para argumentar que, justamente, as queimadas são causadas *por ação humana* e, portanto, não ocorreriam naturalmente, seu interlocutor já está um passo adiante ao colocar a culpa das queimadas em uma famosa estrela de Hollywood. Como você desmente *isso*? Bom, esse é o ponto — engajar-se racionalmente com uma pessoa que argumenta dessa maneira significa sempre gastar mais energia do que ela para desmenti-la e, por isso, talvez não seja a melhor resposta. Mas qual de fato seria? Eu admito certo pessimismo aqui: não acredito que exista uma maneira eficaz de lidar racionalmente com uma pessoa que se porta dessa maneira — e, ao mesmo tempo, abdicar da racionalidade para jogar o seu jogo, no campo dela, por assim dizer, pode ser eficaz mas certamente não seria elegante.

Falhas de racionalidade, vieses e heurísticas

Nós estamos trabalhando até aqui com a ideia de que a lógica é uma disciplina normativa, o que significa que ela estabelece as regras da boa argumentação. E, é claro, nós nem sempre argumentamos logicamente, como quando cometemos falácias. Lembremos que o ponto de partida para entendermos como a lógica opera é a ideia de que temos uma intuição, uma disposição confiável, para reconhecer bons argumentos. Será mesmo que essa intuição é confiável, que ela acerta na maior parte dos casos? Aqui vai um exercício para você. Imagine que você tem diante de si as quatro cartas como na figura abaixo:

Cada uma das cartas tem um numeral em um lado (ímpar ou par) e uma letra de outro (vogal ou consoante). Agora imagine a seguinte regra: *se uma carta tem um número par em uma face, então na outra face ela tem uma vogal*. Imagine que você tem que testar a verdade dessa regra e que você deve fazer isso virando o mínimo de cartas possíveis. Quais cartas você viraria?

Você provavelmente pensou em primeiro virar a carta com o número 2. Vamos imaginar que há uma vogal do outro lado. Até aí, tudo bem. Mas — como você deve ter notado — isso não testa suficientemente a regra acima. Então, você também provavelmente pensou em virar a carta com a letra A. Essa minha estimativa do que você provavelmente faria não é aleatória: o experimento conduzido nos anos 60 por Peter Wason, um psicólogo cognitivo inglês, mostrou que cerca de 90% das pessoas escolhem errado a segunda carta, selecionando a carta da letra A. Por quê? Pense novamente na regra que apresentamos: *se uma carta tem um número par em uma face, então do outro lado ela tem uma vogal*. Ou seja, resumidamente: *se par, então vogal*. Vamos supor que você tenha descoberto um número ímpar no verso da carta A. O problema é que essa descoberta não é relevante para a verdade ou falsidade da regra, porque a regra afirma apenas que, *se par, então vogal*. Ou seja, é perfeitamente possível que ela tenha uma letra vogal em um verso *sem ter* um número par no outro. Por isso, a carta que você deveria ter virado é a D, porque, se você encontrasse um número par no verso dessa carta (consoante), então a regra teria sido mostrada falsa. A regra usada nesse exemplo trata-se do que chamamos em lógica de *condicional material*. A ideia é que, para duas proposições quaisquer *P* e *Q*, 'se *P* então *Q*' é falsa apenas se *P* for ver-

dadeira e se *Q* for falsa — em todos os outros casos, inclusive naqueles em que *P* é falsa e *Q* é verdadeira, a condicional é verdadeira. O modo tradicional de esquematizar a condicional material é a seguinte tabela de verdade ('V' e 'F' abaixo indicam verdadeiro e falso):

P	***Q***	**Se *P*, então *Q***
V	V	V
V	F	F
F	V	V
F	F	V

A descoberta de Wason foi impressionante, porque ela mostra que a maioria das pessoas — inclusive pessoas familiarizadas com lógica — não raciocinam conforme uma regra básica da lógica clássica. Essa descoberta, é claro, requer uma explicação: por que, afinal de contas, desviamos tanto da lógica? Seríamos nós amplamente ilógicos, irracionais? Seria um problema interpretativo da nossa parte quando nos confrontamos com a tarefa de Wason? Ou seria talvez algum defeito na concepção do estudo original, algo que interfira negativamente na nossa capacidade inferencial? Será o caso de revisarmos a própria lógica a partir de uma descoberta empírica, como advogariam os empiristas radicais? Será que o problema talvez esteja com a própria interpretação material da condicional? Se esse for o caso, o filósofo estoico Crísipo estaria vindicado, pois foi ele um dos primeiros defensores de uma interpretação do conectivo 'se… então' diferente da material condicional, já no século III a.C.. De qualquer modo, mais de 300 estudos inspirados na tarefa de seleção de Wason, como é conhecida, foram realizados desde a publicação do artigo original, e alguns obtiveram descobertas interessantes. Um desses casos é o estudo conduzido por Leda Cosmides e John Tooby no início dos anos 90. Seu experimento consistia em substituir a regra original do exemplo de Wason pela seguinte regra: *se uma pessoa está bebendo*

uma bebida alcóolica, ela tem mais de 18 anos. Assim, os pesquisadores também substituíram as cartas por, digamos, '16', '25' (numerais indicando idade), uma imagem de cerveja e uma imagem de um suco. Quanto ao resto, a tarefa era a mesma: virar o mínimo de cartas para descobrir se a regra é falsa. Cosmides e Tooby notaram que há uma maior probabilidade de as pessoas fazerem a escolha correta para realizar essa tarefa: maioria das pessoas vira as cartas contendo o numeral 16 e a imagem de cerveja. Note que se trata do *mesmo raciocínio* que na tarefa original de Wason. Isso sugere que talvez as pessoas tenham mais facilidade para realizar raciocínios em contextos familiares, como os que envolvem normas sociais, do que em contextos mais abstratos, como os que envolvem numerais e vogais sem qualquer relação previamente estabelecida entre si.

Desde meados do século passado, os estudos em psicologia cognitiva ampliaram substancialmente o entendimento sobre como de fato inferimos. Hoje sabemos que nós geralmente usamos *heurísticas*, procedimentos que consomem pouca energia no processamento cognitivo (ao contrário de raciocínios que realizamos conscientemente) e que servem basicamente de *atalhos cognitivos*. Um caso bem estudado é a heurística que nos permite prever a trajetória de um objeto em movimento parabólico. Somos perfeitamente capazes de prever onde cai um objeto arremessado na nossa direção, mas como? Uma hipótese é de que realizamos inferências sofisticadas que envolvem muitas variáveis, como a velocidade do arremesso, a força gravitacional, a ação do vento, etc. E que, além disso, enquanto o objeto está em curso, nós atualizamos essas inferências com novas informações. Segundo essa hipótese, nossos cérebros seriam como supercomputadores que processam informações de modo muito sofisticado, mas abaixo do limiar da consciência — porque nós não temos controle sobre esse processamento informacional. Mas, é claro, há uma maneira muito mais simples de explicar como somos capazes de prever a trajetória de um objeto, a saber, seguindo a chamada *heurística do olhar*: para rastrear um objeto, é preciso observá-lo e manter constante o ângulo, movimentando-se para frente ou para trás. Essa heurísti-

ca dispensa o enorme processamento cognitivo que seria necessário de acordo com a outra hipótese. Uma das principais evidências a seu favor é o fato de que jogadores profissionais de basebol (pessoas com excelente capacidade preditiva sobre a trajetória de objetos) não teriam uma taxa de acerto muito maior do que leigos quando são solicitados a prever o ponto de chegada de um objeto *sem poder ajustar o ângulo de observação*. Esse é o tipo de descoberta que sugere que nosso sistema cognitivo não é como um computador superpotente. Nossa cognição, inclusive, envolve movimentos do corpo todo, não apenas o processamento de informação que ocorre no cérebro.

Além disso, hoje temos uma compreensiva taxonomia dos nossos *vieses*, que também são considerado atalhos cognitivos, embora geralmente não sejam confiáveis (ao contrário de heurísticas — mas, admitidamente, diferença aqui é um pouco difusa). O consenso é que vieses devem ser evitados, seja na prática científica, seja no nosso dia-a-dia. Mas nem sempre é fácil combatê-los, porque vieses são hábitos muito bem arraigados e muitas vezes incorremos em erros desse tipo sem sequer nos darmos conta. Um dos vieses mais populares é o *viés da confirmação*: a tendência de procurar evidências que confirmam as crenças que a pessoa já tem, como quando ela já se decidiu por um político como o seu candidato e evita a todo custo informações que podem prejudicá-lo na sua perspectiva. Isso, por sua vez, geralmente leva as pessoas a formarem *câmaras de eco*, que, como o nome sugere, são ambientes em que as mesmas opiniões são repetidas. Nesses ambientes há pouco ou nenhum desafio crítico às opiniões dos participantes, então é fácil reforçar opiniões seletivamente com base nas informações lá veiculadas.

Outro conhecido viés é o *viés de retrospectiva*: a tendência de pensar que eventos passados eram previsíveis no momento em que ocorreram, sem que nenhuma informação de fato permitisse atestar a sua previsibilidade. Um estudo conduzido pelos psicólogos Incheol Choi e Richard Nisbett mostrou que orientais (especialmente do leste asiático) e ocidentais têm diferentes tendências para cometer o viés da retrospectiva. O estudo foi conduzido da seguinte maneira: sul-co-

reanos e norte-americanos foram expostos a uma história sobre um bom samaritano que estava atrasado para um compromisso e que encontrou alguém com dor no caminho. Ambos os grupos atribuíram inicialmente probabilidades semelhantes (cerca de 80%) de que o bom samaritano ajudaria a pessoa em necessidade. Em uma condição do experimento, foi-lhes relevado que o samaritano na verdade *não ajudou* a pessoa machucada, e então os sujeitos do experimento foram questionados sobre *o que eles acreditariam caso não soubessem disso*. Americanos em geral manifestaram surpresa e mantiveram a mesma probabilidade de que o bom samaritano ajudaria, enquanto coreanos disseram que era igualmente plausível que o bom samaritano ajudasse quanto que não ajudasse (cerca de 50%). É claro, a constatação dessa diferença é por si só interessante, mas o trabalho ainda mais interessante é explicar o porquê dessa divergência.

Há também o famoso *viés das mãos quentes* (às vezes chamado de falácia das mãos quentes ou da mão santa), em que pessoas acreditam que, a partir de um acerto em uma tentativa que depende da habilidade de um agente, é mais provável que o agente acerte novamente no futuro. Esse é um erro porque cada tentativa é independente das anteriores e, portanto, mantendo a habilidade do agente constante, um acerto não confere maior probabilidade de um acerto posterior. Por outro lado, temos o *viés do jogador* (ou falácia do jogador), a tendência de pensar que fracassos sucessivos fazem com que uma próxima tentativa tenha maiores chances de sucesso, mesmo que cada tentativa seja independente das anteriores. O mesmo grupo de pesquisadores trabalhando com Nisbett e Choi também descobriu que pessoas orientais e ocidentais têm diferentes sensibilidades a cada um desses vieses, sendo o viés das mãos quentes mais comum entre ocidentais, enquanto o viés do jogador é mais comum entre orientais.

Outro viés bem documentado é o *efeito de Dunning–Krueger*, que é a tendência, por pessoas com baixo conhecimento ou entendimento sobre um assunto, de superestimar seu próprio estado cognitivo. Um caso familiar é das pessoas que, nunca antes tendo estudado epidemiologia (talvez nunca antes tendo ouvido falar do assunto),

informaram-se por um punhado de vídeos na internet e passaram a considerar-se especialistas sobre pandemias, um assunto obviamente complexo e que não se pode dominar assistindo 5 minutos de vídeos amadores. Se você acompanhou a crise do COVID–19, você viu a emergência de centenas de pessoas sendo vítimas desse viés, propondo "soluções" simples que os especialistas supostamente deixaram de ver porque não estariam pensando o bastante sobre o assunto (por surreal que isso pareça). Em resumo, é a ideia de que um incompetente é incompetente demais para dar-se conta da sua própria incompetência.

Considerações finais

Neste capítulo, examinamos alguns conceitos fundamentais para entender apenas o mínimo de lógica: *verdade*, *argumento*, *formalização*, *validade*, *solidez*, *falácias*. Além disso, vimos um estudo clássico de psicologia cognitiva que constata que a maioria das pessoas não infere de acordo com as normas da lógica. Muitas descobertas dessa natureza têm sido feitas, e podemos interpretá-las como desvelando o seguinte problema genuinamente filosófico: mesmo concedendo que a lógica não se propõe uma disciplina descritiva — de modo que não é exatamente surpreendente que nossas inferências ordinárias não sigam os preceitos estabelecidos pela lógica — o fato de que há divergências abrangentes entre o modo como inferimos e o modo como deveríamos fazê-lo sugere que talvez o problema não seja conosco, e sim com a própria lógica. Qual é, afinal de contas, a fonte da normatividade da lógica? Ela deveria, como disse Wittgenstein uma vez, cuidar de si mesma? Ou seria a lógica, em última análise, uma abstração feita a partir das nossas inferências ordinárias e que, em casos de divergências profundas entre a lógica e nossas práticas inferenciais, deveria a lógica ser revista sob a luz do modo como de fato inferimos? Essas são questões profundas da filosofia da lógica, e respondê-las escaparia das nossas pretensões neste livro.

Leituras recomendadas

Para uma compreensiva introdução à lógica, a melhor recomendação é sempre Mortari (2001). Sobre noções gerais de argumentação e sobre as noções de validade e solidez, veja os capítulos um e dois de Nolt e Rohatyn (1991). Veja o capítulo sete desse mesmo livro para uma lista de falácias. Recomendo também a consulta ao dicionário de lógica de Hegenberg e Silva (2005). O imenso livro de William e Marta Kneale (1991) é essencial para o entendimento profundo da história da lógica, incluindo discussões sobre silogística, lógica contemporânea e a frequentemente ignorada lógica estoica. Também acerca de lógica estoica, o trabalho de Dinucci e Duarte (2016) é pormenorizado, acessível e claro. Sobre uma avaliação breve da silogística aristotélica, mas muito precisa e contundente, veja Murcho (2019). Sobre teorias da verdade, veja Costa (2005). O livro de Kahneman (2011) é um clássico contemporâneo da psicologia cognitiva, em que o autor discute de modo acessível vieses e heurísticas.

5

Filosofia da Ciência e a crise de racionalidade

Ciência em evidência

Em abril de 2020, pelo menos ⅓ da população mundial encontrou-se em algum estado de isolamento social, seja em distanciamento voluntário ou em *lockdown* imposto pelas autoridades — medidas que foram tomadas na tentativa de conter a pandemia de COVID–19. Além de afetar profundamente os nossos costumes sociais, a disseminação de um vírus capaz de causar o colapso de sistemas de saúde trouxe consigo outra consequência, uma especialmente interessante para a filosofia da ciência: a pandemia colocou discussões sobre o alcance e a relevância da ciência na boca do povo, nos jornais, nas redes sociais. Mais do que isso: faz-se clara uma disputa entre, por um lado, aqueles que tentaram pautar a conduta pública com base nas melhores evidências disponíveis sobre um assunto complicado e, por outro, o diz-que-me-disse e assim-me-parece característico de epidemiologistas e economistas de poltrona com pós-graduação em mídias alternativas.

A ciência está, então, em evidência. Na verdade, não é como se algum dia ela tivesse deixado de ser importante. A ciência é nosso melhor instrumento para promover entendimento da realidade e para intervir com sucesso em um mundo complexo — há quem defenda que a ciência oferece um modelo exemplar de conhecimento, em comparação com o qual devemos avaliar todos os outros supostos conhecimentos. De qualquer modo, a pandemia trouxe consigo

uma roupagem nova para debates filosóficos clássicos sobre ciência: por um lado, temos evidências científicas para crer em determinadas hipóteses e, por outro, temos nossas convicções de que as coisas são assim-e-assado, e frequentemente ambas não se coadunam pacificamente. O que é pior, descobertas científicas são custosas, tanto do ponto de vista material quanto humano: amplos investimentos em pesquisas e em formação de profissionais qualificados são indispensáveis para que possamos chegar a conclusões científicas robustas, o que custa caro e toma tempo. Opinar, por outro lado, é rápido, barato, e pode ser feito sem sair de casa. Além disso, alguns céticos (alguns deles sem dúvida mal-intencionados) questionam a confiabilidade das descobertas científicas com base nos fatos de que cientistas frequentemente divergem entre si e, além disso, de que não é incomum que uma hipótese defendida ontem tenha sido refutada hoje. Para piorar a situação, a linguagem científica é técnica, o que significa que ela não é imediatamente traduzível para leigos, e os métodos científicos tornaram-se refinados a ponto de parecerem incomensuráveis com nossas práticas investigativas de senso comum. A complexidade do linguajar especializado e dos métodos científicos pode, então, gerar uma frustração para quem não tem alguma familiaridade prévia com os fundamentos da ciência. Esse sentimento de exclusão do *mainstream* científico, por sua vez, empurra os cognitivamente menos qualificados às margens da produção de conhecimento e de entendimento sobre o mundo. Especulo — e isso é apenas uma especulação — que desse movimento de, digamos, êxodo epistemológico é que nascem as teorias da conspiração e as comunidades que rejeitam a racionalidade científica. Para os membros desses grupos, conspirar contra a ciência tem o duplo papel de, por um lado, defender as convicções de senso comum (por mais frágeis que sejam) contra as "ameaçadoras" descobertas científicas e de, por outro, criar um sentimento de união entre os excluídos, de promover uma sensação acolhedora de pertença a um grupo de semelhantes. Grupos assim não apenas se opõem a descobertas científicas, eles promovem os seus próprios dogmas como se fossem uma alternativa à ciência,

uma “para-ciência” que teria a vantagem de ser mais inclusiva do que a sua rival institucionalizada, porque não discrimina ideias espúrias. Visto que estupidezes são fáceis de produzir, pois qualquer um pode enunciá-las da poltrona de casa sem nenhum esforço ou compromisso, os movimentos anticientíficos tendem a se ampliar facilmente.

Esse é o caso, por exemplo, dos terraplanistas. Você provavelmente já ouviu falar dessas pessoas. Há alguns anos elas não passavam de uma piada de internet. Hoje elas influenciam políticas públicas, e não vai demorar muito para que seu dogma entre nos currículos escolares, o que eu suspeito que abrirá a válvula para outras aberrações pseudocientíficas. A crença que une os terraplanistas é, como o nome sugere, a ideia abismal de que a Terra é plana. Por que essas pessoas acreditariam nisso, você se pergunta, visto que há registros fotográficos da Terra vista do espaço? Por que acreditar nesse absurdo quando, já no século III a.C., o astrônomo e matemático grego Eratóstenes foi capaz de constatar a circunferência da Terra com uma margem de erro de apenas 300km? Vamos imaginar que a Terra fosse de fato plana: por que ela haveria de ser diferente dos outros corpos celestes? E por que plana, e não, digamos, em formato de Doritos? A resposta a todas essas perguntas é: por nenhuma razão em especial senão pelo desprezo pelo que a ciência representa. É um ato de rebeldia, não diferente da familiar rebeldia adolescente, mas nesse caso ela é frequentemente tardia — o que nos leva a crer que essas pessoas ainda não atingiram a maturidade intelectual, ou talvez que voluntariamente escolheram evitá-la.

O terraplanismo é uma ideia ridícula, no sentido literal de que convida ao riso, mas ela anda de mãos dadas com ideias legitimamente perigosas, como o movimento antivacina. Esse movimento nasceu com um artigo fraudulento publicado em 1998 por um então médico que supostamente havia constatado uma correlação entre vacinação e autismo em crianças, sugerindo que vacinas seriam a causa do autismo. É claro, estudos sérios comprovaram que não há nenhuma evidência sequer em favor dessa relação causal, e o charlatão (cujo nome não merece menção) foi amplamente desacreditado

pela comunidade científica e acabou perdendo o seu direito de praticar medicina. Mas o mal já estava feito, pois o movimento antivacina tornou-se viral, literal e metaforicamente. Rapidamente se espalhou pelo mundo e até hoje é diretamente responsável pela morte de centenas de crianças todos os anos. Se isso não fosse o suficiente, a onda antivacina solapa os pressupostos da chamada *imunidade de rebanho*, que ocorre quando grande parte da população está imunizada a uma doença e que, portanto, diminui as chances de transmissão para aqueles que não podem ser imunizados por razões muito específicas e raras. O fato de que há pais que não vacinam seus filhos contra a rubéola, por exemplo, fez com que o número de infecções em 2019 nos Estados Unidos disparasse, tornando-se o maior desde 1992 — o que é especialmente preocupante porque a doença havia sido considerada erradicada no ano 2000. Não há dúvidas de que denunciar a fraude do "estudo" original e excluir o responsável da comunidade científica foram as atitudes corretas, mas elas também tiveram um efeito psicológico curioso (e preocupante) entre os proponentes dessa ideia nefasta. O sentimento de exclusão dos que se deixaram persuadir pelo embuste foi intensificado, alimentando a suspeita absurda de que as grandes organizações de saúde fariam de tudo para "abafar" uma verdade à qual apenas os antivacina teriam acesso. Essas pessoas veem-se como mártires enquanto seus filhos definham de doenças perfeitamente evitáveis.

Certos posicionamentos que presenciamos com o avanço do novo coronavírus não foram muito diferentes das atitudes anticientíficas disseminadas entre os grupos terraplanista e antivacina. Em menos de 3 meses de pandemia, foi veiculado o factoide de que o vírus havia sido criado em um laboratório chinês com o intento de enfraquecer as economias mundiais e com isso criar oportunidades para investidores chineses. É claro, pôde-se constatar pelo estudo criterioso do RNA do vírus que não se trata de um vírus manufaturado. Mas essa descoberta foi recebida por alguns com *ainda* maiores suspeitas, o que gerou a infeliz reação de dúvida quanto às reais intenções das instituições científicas por ficarem "do lado da China". E quanto às pes-

soas que acreditavam que pandemia é uma farsa, que não há mortes pelo vírus e que tudo teria sido uma armação com fins políticos? O que se poderia fazer para mudar a opinião de uma pessoa que não se deixa convencer pelo que dizem os especialistas em um assunto, ou até mesmo o bom senso? Confrontar uma pessoa dessas com novas evidências baseadas no que dizem os cientistas não funcionaria, porque, paradoxalmente, os mais qualificados tornaram-se os menos confiáveis na sua visão distorcida. Isso nos mostra um problema filosófico e social legítimo, um sintoma de uma crise de racionalidade que transcende o fenômeno específico da pandemia de 2020. Como convencer quem não compartilha dos mesmos ideais de racionalidade que nós? Argumentar contra uma pessoa que não reconhece o valor racional da argumentação já é em certo sentido uma derrota: como diz a piada, é como tentar jogar xadrez contra um pombo — ele derruba as peças, defeca no tabuleiro e empertiga-se como se tivesse ganhado o jogo. Como poderíamos *debater* com uma pessoa dessas? Parece vão esperar de braços cruzados enquanto essa pessoa dissemina ideias que põem a si e aos outros em risco, e parece ainda mais inútil torcer para que ela eventualmente tenha um momento de iluminação e revise suas opiniões originais. Talvez a estratégia mais producente seja minimizar o impacto público do seu posicionamento, pois, embora seja quase impossível convencê-la, talvez ainda seja possível evitar que ela influencie os demais. O problema é que não é claro como isso poderia ser feito. Devemos ignorá-la? Isso, a meu ver, só facilitaria a percepção errônea de que a sua posição é incontestada. Devemos tomar atitudes menos diplomáticas e ridicularizar suas crenças, expondo-a aos seus pares como uma figura que não merece ser levada a sério? Isso obviamente não é muito elegante, e talvez não seja eficiente.

A pandemia também colocou as evidências em evidência, com o perdão da redundância. A busca por uma resposta eficaz ao coronavírus rapidamente gerou um par de estudos relâmpago que supostamente teriam mostrado que alguns compostos usados no tratamento de outras doenças poderiam ser promissores no combate à infecção

causada pelo novo coronavírus. Os resultados pareciam empolgantes, o que motivou campanhas enfáticas pelo fim do isolamento social sob a suposição de que as enfermidades relativas ao novo coronavírus poderiam ser combatidas por uma medicação que já era produzida em larga escala e que já era aplicada com sucesso no tratamento de outras doenças. Melhor ainda do que uma vacina, alguns prontamente pensaram (talvez por pensarem também que vacinas causam autismo). O problema? Os estudos foram recebidos pela comunidade científica como amplamente inconclusivos. Ou seja, não haveria evidências fortes o suficiente em favor da eficácia daquelas substâncias. Um dos estudos havia mostrado que o medicamento mata o vírus in vitro — sim, pode ser verdade, mas um tiro de revolver também mataria qualquer coisa in vitro. O ponto é que um resultado obtido in vitro não pode ser facilmente extrapolado para o corpo humano porque esses são ambientes muito diferentes. O outro estudo preliminar consistiu na aplicação de outro medicamento em alguns pacientes com o vírus. O estudo envolveu 30 pessoas divididas em dois grupos, em que as pessoas de um grupo receberam esse tratamento e as do outro serviram de controle (isto é, não receberam a mesma medicação). Além da amostragem pequena, a taxa de melhora foi basicamente idêntica em ambos os grupos. Ou seja, as evidências não permitiram concluir pela eficácia do tratamento. Sabe-se também que esses compostos, se mal administrados, podem levar a graves problemas (como dificuldades cardíacas, cegueira e surdez) e inclusive ao óbito. De qualquer modo, é possível que o avanço das pesquisas mostre o sucesso dessas medicações no tratamento da COVID–19? É claro que é — e eu acredito que maioria das pessoas *espera* que um tratamento eficaz (se não esses, pelo menos outro), seja descoberto em breve. Mas *esperar* que exista uma cura não deveria ser confundido com *acreditar* que existe uma cura, nem ser usado de pretexto para vendê-la como uma solução. Fundamentalmente, o ponto é que, dadas as melhores evidências das quais dispostos presentemente, é profundamente imprudente apresentar como panaceias medicamentos cuja eficácia não está comprovada.

No horizonte de todas essas discussões estão algumas das questões que apresentaremos aqui, como: será possível atestar a verdade de uma hipótese científica? O que difere ciência genuína de pseudociência? Qual é, afinal de contas, o método científico, se é que existe um *único* método? Como ocorre a prática científica, e qual o papel da comunidade científica em escolher o que é ou não ciência? Pode a ciência conquistar conhecimento, se eventualmente todas (ou quase todas) teorias científicas são rejeitadas?

Conhecimento ordinário e conhecimento científico

Pense por um momento no que você sabe. Com alguns momentos de atenção, você é capaz de fazer uma ampla lista. Provavelmente, alguns dos conhecimentos que você listou são sobre objetos, pessoas ou lugares, alguns são sobre como fazer certas ações, alguns são sobre estados de coisa, isto é, conhecimentos de que as coisas são assim-e-assado. Parece haver, pelo menos do ponto de vista gramatical, uma diferença entre conhecer a cidade de Salvador e saber que Salvador é uma cidade. Do mesmo modo, parece haver uma diferença entre saber que o que vai na receita de risoto e saber fazer um risoto. Eu discuti extensivamente essas e outras questões no meu livro *Epistemologia: uma Introdução Elementar* (2018), e eu sugiro a leitura desse material se você estiver interessado(a) em questões de epistemologia tradicional, como, por exemplo, a diferença entre os tipos de conhecimento, ou a possibilidade de conhecimento sobre o mundo exterior e sobre o que pensam outras pessoas. Neste capítulo, nossa estratégia é outra, pois trabalharemos com filosofia da ciência, um assunto que havia ficado de fora do outro livro. Não há, no entanto, necessidade da leitura prévia daquele material para as nossas presentes discussões.

Conhecer envolve, de modo muito geral, *sucesso* ou *êxito*. Você não pode saber que a Lua é feita de queijo porque ela *não é*. Você poderia acreditar que ela é feita de queijo, e você estaria errado, teria fracassado na sua crença. Quando tratamos desse tipo de conheci-

mento, o *saber que* algo é assim-e-assado, o sucesso em questão é a *verdade*. Ou seja, você não pode saber que alguma coisa é de certo modo se aquela coisa não é daquele modo. O que é interessante é que essa mesma regra vale tanto para o conhecimento ordinário ou prosaico, aquele que adquirimos no nosso dia-a-dia, seja pelas nossas percepções, inferências, ou pelo testemunho de outrem; quanto para o conhecimento científico, que depende — além de percepções, inferências e testemunho de especialistas — de uma arcabouço técnico e institucional muito mais sofisticado do que o nosso conhecimento ordinário. Alguns aspectos desse arcabouço são a replicabilidade de experimentos científicos (experimentos bem conduzidos podem ser repetidos com taxa de sucesso semelhante) e a avaliação por pares (em que outros especialistas na área avaliam, idealmente sem conhecer os autores do estudo sob avaliação, se os procedimentos empregados foram adequados). Apesar das diferenças, se forcamos no conceito de sucesso envolvido em ambos os casos, veremos que há algo em comum entre os dois tipos de conhecimento, o conhecimento científico e o conhecimento ordinário: ambos implicam sucesso. Se descobrirmos que uma hipótese científica é falsa, mesmo tendo sido previamente aceita por toda a comunidade científica — como foi para a hipótese geocêntrica antes da revolução copernicana, por exemplo — então é porque essa hipótese não se tratava de *conhecimento* afinal de contas.

Além disso, para ambos os casos, o sucesso não pode ser atingido por acaso. Assim como um chute certeiro em uma prova objetiva pode render-lhe uma questão correta, mas que, por tratar-se de um chute, não demonstra o seu conhecimento sobre a questão; uma cientista que é levada a crer na hipótese verdadeira pelas razões erradas e evidências inadequadas também não tem aquele conhecimento em particular. Por conta disso, os procedimentos que são usados na investigação científica são especialmente importantes. O modo como cientistas defendem suas conclusões é o chamado *método científico*. Dito isso, a pergunta que nos interessa agora é: qual o método,

ou quais os métodos, que empregamos na obtenção do conhecimento científico?

Indução e abdução

Esta é uma história real: nos primeiros dias da quarentena, um vizinho do prédio em frente decidiu tocar, às 18h, a música *Ave Maria* cantada pelo tenor Luciano Pavarotti em um grande alto-falante voltado para a rua. No começo, foi uma agradável surpresa, que talvez tenha sido acompanhada da intenção de transmitir uma mensagem de paz e de união. O evento inicial, é claro, gerou aplausos de muitos vizinhos. Esse ato tornou-se uma regularidade pontual nos dias seguintes, de modo que, quando começa a cair a noite, já formamos a expectativa de ouvir a música às 18h em volume ensurdecedor. É claro, agora nossa expectativa é acompanhada por profundo aborrecimento, não pela música, mas pelo fato de que a intenção original parece ter sido substituída por outra expectativa: o vizinho parece repetir seu show tão somente para receber as poucas palmas das senhorinhas emocionadas, e nós somos obrigados a parar tudo que estamos fazendo e ouvi-lo.

Juízos de valor à parte, o essencial desse exemplo é o seguinte: um mesmo acontecimento ocorreu regularmente no passado, de modo que é possível inferir que o mesmo evento ocorrerá novamente no futuro. Nesse exemplo especificamente, a partir das observações passadas de que a música foi tocada no dia 1, no dia 2, no dia 3, etc., inferimos que ela será tocada amanhã. Essa inferência é uma inferência *indutiva* (o que não tem nada a ver com "induzir ao erro") e, diferentemente das inferências dedutivas, que comentamos no capítulo anterior, ela é não-monotônica. Uma inferência é monotônica se, se ela é válida, não importa quais novas premissas são acrescentadas ao conjunto inicial de premissas, ela continua sendo válida. Pense, por exemplo, no seguinte argumento dedutivo: Se todo homem é mortal, e Sócrates é homem, então Sócrates é mortal. Trata-se de um argumento válido e, mesmo que acrescentássemos que a Lua é

feita de queijo entre as premissas, ele continuaria sendo válido. Inferências indutivas não funcionam desse mesmo modo, pois o acréscimo de uma nova evidência ao que até então era uma boa inferência indutiva pode ter o efeito de desautorizar a conclusão. Basteria, no nosso exemplo inicial, um bendito dia em que o vizinho tivesse um lampejo de lucidez e de consideração pelo bem-estar alheio, de modo a decidir não tocar a música. Assim, se no dia *x* a música não é tocada, mesmo tendo sido tocada em todos os outros dias da quarentena, já não temos mais a mesma força para concluir que a música será tocada amanhã.

Diferentemente de inferências dedutivas, inferências indutivas não almejam a *validade*. Quando inferimos dedutivamente, por exemplo, que, se Ostara está sobre a mesa, então Klaus está no quarto, e Ostara está sobre a mesa, logo, Klaus está no quarto, pretendemos que a conclusão se siga das premissas. Isso ocorre até mesmo quando falhamos na nossa inferência. Se eu cometesse uma falácia de afirmação do consequente, eu ainda pretenderia — salvo em situações retóricas — que a conclusão se seguisse das premissas. Mas o mesmo não é o caso para inferências indutivas. Pense no seguinte exemplo: você observou um cisne branco, dois cisnes brancos, três... cem... mil cisnes brancos, e dessas observações você infere que *todos os cisnes são brancos*. Essa conclusão diz respeito a *todos* os cisnes, isto é, ela diz de todos eles que eles são brancos, mesmo aqueles que existiram muito antes de você nascer e aqueles que nascerão muito depois de você ter virado pó. Desse modo, é perfeitamente possível que, se interpretarmos cada observação particular como uma premissa em um longo argumento, as premissas sejam verdadeiras sem que a conclusão também seja. Isto é, mesmo que todos os cisnes que você observou tenham sido de fato brancos, é possível que exista pelo menos um cisne não branco ainda não observado. Mas isso não faz dessa inferência inválida, *porque ela não se pretende válida*! Seria um erro categorial caracterizá-la como uma inferência "indutivamente inválida", porque inferências indutivas pretendem conferir força ou plausibilidade à conclusão, não pretendem derivá-la das premissas.

Nesse caso, as observações de mil cisnes brancos parecem conferir alguma força à conclusão de que todos os cisnes são brancos. Mas isso é claro depende de outros fatores, como quão ampla e aleatória é a sua amostragem. Digamos que que você observou mil cisnes brancos na sua cidade, mas que há dados que sugerem que existam atualmente pelo menos 500 mil cisnes espalhados pelo mundo. A sua amostragem, nesse caso, foi pequena e possivelmente não representativa. Se você tivesse observado mil cisnes brancos selecionados aleatoriamente em várias cidades diferentes, sua conclusão teria maior suporte indutivo.

Além de estar presente no nosso dia-a-dia (como no caso do vizinho inoportuno), a indução parece ser um procedimento científico por excelência. Quando uma cientista pretende constatar, por exemplo, a verdade da hipótese de que a doença *D* pode ser tratada com sucesso através da administração do medicamento *M*, ela o faz através de uma série de observações que conferem maior ou menor força à hipótese em questão. É claro, essa é uma simplificação muito grosseira, mas ela captura de maneira muito geral a estrutura de uma hipótese científica. Importantemente, embora não pareça, a hipótese abrevia uma generalidade, pois ela não está falando do tratamento da doença de um paciente em particular, mas de (se possível) todos os pacientes com aquela doença. Assim, talvez devamos interpretá-la deste modo: *todos os casos da doença D podem ser tratados com sucesso através da administração do medicamento M*. É claro, no mundo real, as coisas são bastante mais complicadas, e é possível que existam variáveis que condicionem a hipótese e delimitem a sua aplicação. Por exemplo, um paciente afetado pela doença *D*, mas com histórico de outra doença *D**, talvez não possa ser tratado com *M*, ou com a mesma dosagem, etc. Desse modo, talvez a hipótese deva ser revista: *todos os casos da doença D, exceto aqueles em que há histórico de D*, podem ser tratados com o medicamento M*. Qualquer que seja o caso, a indução parece desempenhar um papel importante, porque é ela que permite partir do observado para o ainda inobservado — que nesse caso seria o sucesso da administração do tratamento para pacientes futuros.

Há também inferências analógicas, que são não-monotônicas, e que podem ser consideradas inferências indutivas. É comum na prática científica inferir que certas observações feitas em um domínio restrito possam ocorrer em outro domínio a partir da suposição de que há algo análogo em ambos os casos. Mais especificamente, considere a administração de medicamentos em certos animais não-humanos para a observação das reações possíveis que ocorrem nesses casos. Verificadas essas reações, e com a suposição de que há analogias biológicas importantes entre esses animais e seres humanos, é possível inferir analogicamente que seres humanos terão as mesmas reações, ou reações semelhantes. Novamente, novas evidências podem derrogar a conclusão que havia sido estabelecida, como a descoberta de uma inadequação da analogia entre os animais não-humanos e nós no que diz respeito a um traço biológico importante. Inferências analógicas são como inferências indutivas enumerativas (isto é, inferências que dependem da enumeração de casos) porque, além de derrogáveis pela nova evidência, partem do observado para o não-observado.

Tamanha parece ser a importância da indução que o positivismo do século XIX defendia que (1) há um único método para a obtenção do conhecimento científico e que (2) esse método consiste na indução, pois, a partir de um conjunto de observações particulares, seria possível inferir a verdade de uma hipótese universal, isto é, um enunciado do tipo *todo A é B*. Hoje nos parece, no entanto, que essas ideias são muito restritivas. Elas excluem de partida, por exemplo, as ciências humanas. Imagine que você é um sociólogo que busca estudar os efeitos da quarentena nas nossas práticas sociais prolongadas. Plausivelmente, quando a quarentena acabar, as coisas voltarão aos poucos ao normal. Será mesmo? Essa é uma questão empírica, e é necessário vê-la sob várias perspectivas diferentes. Mas como poderíamos conduzir um estudo que permitisse inferir indutivamente a verdade da hipótese, digamos, de que as nossas práticas sociais prolongadas voltarão ao normal? Em primeiro lugar, essa não é uma hipótese universal. Ela diz respeito às *nossas* práticas sociais, não às prá-

ticas sociais de quaisquer pessoas ao redor do mundo. Ainda assim, talvez possamos reescrever a hipótese da seguinte maneira: *todas as nossas práticas* etc. Mas, a principal dificuldade é que não é possível, nem eticamente desejável, conduzir um experimento em ambiente controlado que permita inferir indutivamente qualquer coisa sobre as nossas práticas. Não seria razoável, no ambiente da pesquisa, controlar e observar um grupo de pessoas e inferir algo sobre o nosso comportamento social a longo prazo, pois, além dos problemas éticos nessa metodologia, há uma pletora de variáveis que interferem no comportamento humano, de modo que qualquer inferência para hipóteses *universais* estaria prejudicada.

Além da indução enumerativa e da analogia, há outro procedimento inferencial não-monotônico que é crucial para as práticas científicas, e ele não está contemplado na antiga visão positivista. Trata-se da inferência à melhor explicação, ou ainda, *abdução* — que nada tem a ver com alienígenas obcecados por sondas. Com efeito, o famoso "historiador" de cabelo desgrenhado que pretende explicar quaisquer fenômenos perfeitamente mundanos pela interferência de alienígenas serve-nos de exemplos do que seria uma inferência para a *pior* explicação. Se pirâmides foram construídas por todo o mundo antes da invenção de instrumentos modernos de engenharia, a melhor explicação é que o formato piramidal (base larga e um estreitamento em direção ao cume) permite mais facilmente empilhar materiais pesados. Nada disso requer a suposição de vida inteligente fora da Terra, que, apesar da tecnologia sofisticada o suficiente para viagens interplanetárias, erigiria monumentos relativamente simples. Isso pode parecer piada, não mais do que um comentário sobre um meme, mas o fato de que explicações infundadas e facilmente refutadas têm alta capilaridade é um sintoma da crise de racionalidade pela qual passamos.

Para entender a abdução mais claramente, pense no seguinte exemplo. Você está em casa sentado no sofá. Mais cedo nesta semana, você fez uma compra pela internet que deve ser entregue por esses dias. A campainha toca, e você não está esperando ninguém. Sem

se levantar do sofá, você infere que deve ser o entregador com a sua encomenda na porta. O que ocorre nesse caso é que você inferiu que a melhor explicação para a campainha ter tocado é a chegada da sua entrega. Ou seja, de um conjunto de informações das quais você já dispunha, você inferiu algo novo que ampliou o seu conhecimento inicial. Essa inferência é, assim como inferências indutivas, sensível ao acréscimo de novas evidências que podem cancelar sua conclusão. Pode ser o caso que seu vizinho tenha tocado a campainha para trazer algo que você emprestou, por exemplo. Se você tivesse acesso a esse fato, essa nova evidência cancelaria sua explicação inicial. Apesar das semelhanças, sua inferência não foi indutiva, porque ela não almeja uma generalidade, pois você quer explicar apenas qual a melhor explicação daquele evento específico, não de todas as vezes em que a campainha toca.

Abduções abundam na prática científica, com o perdão da aliteração. Frequentemente a observação recorrente de um tipo de evento particular é acompanhada da melhor explicação pelo que pode tê-lo causado. Imagine que você é um estudante de química que mistura sódio na sua forma sólida com água. Conhecidamente, isso resulta em uma explosão. A melhor explicação desse fenômeno parece ser que a reação libera outros compostos gasosos e muito calor, de modo a gerar uma combustão. Mas note que pode acontecer algo completamente contrário à sua expectativa: digamos que, em uma experimentação, não houve explosão. Agora a tarefa é buscar a melhor explicação para esse novo evento. Talvez haja alguma variável ainda não contabilizada, talvez você sem querer colocou um cubo de açúcar, não de sódio, na água. É claro, na prática científica, indução e abdução ocorrem concomitantemente. Por exemplo: a observação reiterada da combustão resultante da mistura de sódio sólido com água permite inferir que isso sempre ocorrerá dessa maneira em condições apropriadas, o que por sua vez requer a inferência para a melhor explicação desse fenômeno.

Finalmente, note que, se inferências não-monotônicas como a indução enumerativa e analógica, bem como a inferência à melhor

explicação, forem realmente o meio através do qual se pode conferir graus de plausibilidade a hipóteses científicas, e visto que essas inferências são abertas a novas informações que podem cancelar as suas conclusões, é fácil de explicar porque a opinião de cientistas está sujeita à mudança. Novas observações podem exigir uma mudança de opinião, e isso está perfeitamente de acordo com uma boa conduta epistêmica, seja na ciência, seja no dia-a-dia. Recentemente, o biólogo e divulgador científico Átila Iamarino foi acusado de alarmismo por aplicar um modelo (de modo muito simples, um modelo é uma espécie de simplificação que usa certas informações para tornar um domínio de investigação tratável de uma perspectiva limitada) adaptado do Imperial College de Londres para prever quantos seriam os mortos pelo novo coronavírus no Brasil. Sua estimativa, sem levar em conta muitas variáveis, foi de cerca de 1 milhão de mortos até agosto de 2020. No entanto, o número de mortos está (enquanto escrevo) abaixo do que uma curva exponencial indicaria a ponto de chegar a 1 milhão de mortos em agosto. Iamarino foi alarmista? É claro que não, porque o modelo que ele adaptou fez o cálculo para um cenário sem nenhuma medida de isolamento. O número de óbitos mais baixo do que o esperado dá-se em razão de que os estados, malgrado o desejo fúnebre de alguns, aplicaram as normas de isolamento social em tempo hábil. Esse número também serve de evidência para atualização dos modelos que constatam que o isolamento funciona, mas, principalmente, permite revisar a crença inicial de que haveria 1 milhão de mortos até agosto. Para os cientificamente desavisados, pode parecer que a mudança de opinião é uma falha de racionalidade. No entanto, uma pessoa é mais claramente irracional quando é fundamentalmente incapaz de revisar suas crenças, e isso sim é palpavelmente problemático hoje.

Embora pareça plausível descrever o fundamental da prática científica através de procedimentos indutivos e abdutivos, a indução e a abdução não estão livres de dificuldades profundas.

Enigmas da indução, o antigo e o novo

Passar do observado para o inobservado parece simples e incontroverso. Eu percebi um número incontável de vezes que o Sol nasceu todas as manhãs, e posso facilmente inferir disso que o Sol sempre nasce e que, portanto, também nascerá amanhã. À primeira vista, isso não parece nada misterioso. Mas tem um problema nesse tipo de performance inferencial, um problema que foi primeiro constatado pelo escocês David Hume, o maior filósofo da modernidade, já no século XVIII: a passagem do que observamos para o inobservado depende da suposição do que Hume chamou de *Princípio da Uniformidade da Natureza*, a ideia de que os eventos observados até então continuarão sendo da mesma maneira no futuro. Isso quer dizer que, mesmo tendo observado eventos semelhantes ocorrerem de determinada maneira no passado, como o Sol nascer todos os dias até hoje; a minha inferência para a conclusão de que o Sol nascerá *amanhã* depende da suposição de que a natureza permanecerá constante, como de fato tem permanecido. Mas o que poderia *justificar* a suposição de que as coisas serão do mesmo modo que foram até então?

Essa pergunta leva-nos a um dilema em que nenhuma das duas vias parece promissora. Notemos que o Princípio da Uniformidade da Natureza é uma tese a respeito do mundo, porque se refere à suposta ordem das coisas, a como as coisas são na realidade. Ou seja, trata-se de uma tese *empírica*, uma tese que escapa os domínios do *a priori*. Então não podemos saber *a priori* que as coisas continuarão sendo como foram até então, até porque é perfeitamente possível que elas deixem de ser dessa maneira. Ou seja, negar o referido princípio não implica uma contradição.

Se não é possível justificar *a priori* aquele princípio, resta-nos tentar oferecer uma justificação *a posteriori*, uma justificação por meio da experiência. A pergunta, agora, é: como eu poderia justificar, por meio da experiência, que as coisas continuarão sendo como foram até agora? Talvez a justificação fosse mais ou menos assim: todos os dias eu constatei que, até esse momento, o Princípio da Uniformi-

dade da Natureza estava correto, isto é, que as coisas continuaram ocorrendo de maneira constante, ordenada, previsível, etc. Portanto, o Princípio da Uniformidade da Natureza estará correto no futuro: as coisas continuarão ocorrendo do modo como ocorreram até então. Bem como você pode ter notado, o argumento acima tem um problema: ele pretende *justificar indutivamente* o Princípio da Uniformidade da Natureza — todavia, visto que toda inferência indutiva depende desse princípio, o próprio argumento indutivo que o justifica supõe aquilo que pretende provar. Esse é um caso paradigmático de circularidade, e não são muitos os que aceitariam um argumento circular como convincente.

De certo modo, há ainda uma terceira opção, aquela plausivelmente adotada por Hume, que consiste em rejeitar que haja qualquer necessidade de justificar a indução para além da mera "força do hábito". Estamos acostumados com certas ocorrências terem se repetido de maneira uniforme no passado, e é nosso costume que nos leva a crer que se repetirão do mesmo modo no futuro — estaríamos, segundo Hume, *condicionados*, mas não *justificados* a crer que as coisas continuarão sendo como foram até então. Isso, no entanto, não parece satisfatório: diante de um problema cético, um problema que põe em dúvida a possibilidade do conhecimento obtido pela indução, Hume teria oferecido uma solução cética, segundo a qual o suposto conhecimento indutivo não passaria de um hábito. Isso é menos do que esperamos como resposta adequada ao problema.

Esse problema, que nós chamamos hoje de *antigo enigma* (ou *problema*) da indução, por conta do trabalho do filósofo contemporâneo Nelson Goodman, parte da seguinte constatação: toda inferência indutiva depende da suposição do Princípio da Uniformidade da Natureza, a suposição de que a natureza permanecerá constante. Mas, e esse é o dilema que dá origem ao enigma, esse princípio não pode ser justificado *a priori*, porque diz respeito ao mundo, tampouco pode ser justificado *a posteriori*, porque isso dependeria da indução e seria obviamente circular. Ou seja, embora nossas práticas indutivas pareçam estar presentes em toda nossa vida cognitiva, para a obtenção de

crenças justificadas ordinariamente e para a justificação de hipóteses científicas, a indução ela mesma não está justificada. Note bem que o problema ou enigma da indução *não é* que boas induções podem ter eventualmente conclusões falsas, conclusões que não se seguem das premissas. Isso nunca tirou o sono de ninguém, pois, como vimos, a intenção de argumentar indutivamente é dar suporte ou probabilidade a conclusões, *não é* derivar conclusões das premissas (como em argumentos dedutivos). Em resumo, o problema da indução é que não parece ser possível justificar a nossa confiança na indução ela mesma.

Nelson Goodman cunhou a expressão 'antigo enigma da indução' no seu livro, *Fact, Fiction and Forecast* (em português, *Fato, Ficção e Previsão*), publicado em 1955. Ele usou essa expressão porque tinha em mente um *novo* enigma. Não só isso, ele acreditava que o antigo enigma apresentado por Hume poderia ser facilmente resolvido com algumas considerações sobre as nossas práticas inferenciais. Goodman argumenta que é possível adotar uma justificativa para nossas inferências indutivas, mas essa justificativa não é *a priori*, mas sim dependente da nossa própria experiência indutiva. Ou seja, essa justificativa não nos garante previamente que o princípio humano da uniformidade da natureza é correto. Para mostrar como isso funciona no caso da indução, Goodman começa examinando a justificação para inferências dedutivas. Ele recorre ao procedimento chamado de e*quilíbrio reflexivo*, uma ideia que havia sido apresentada na filosofia política de John Rawls. Segundo esse procedimento, uma inferência dedutiva é aceita como válida se ela concorda com regras que aceitamos — e aceitamos regras como corretas se elas sancionam as inferências que fazemos. O equilíbrio reflexivo, portanto, envolve um ajuste mútuo entre regras aceitas e inferências realizadas, de modo que, se uma inferência viola uma regra que não estamos dispostos a retificar, a inferência é rejeitada, e se uma regra proíbe uma inferência que não estamos dispostos a deixar de realizar, a regra é revisada. O mesmo ocorre no caso de inferências não-dedutivas. Assim sendo, a indução está justificada em virtude das nossas práticas inferenciais

indutivas, do modo como de fato inferimos indutivamente. De um ponto de vista geral, é claro, essa resolução do antigo enigma envolve adotar a via empírica e, portanto, circular de solução do dilema, porque a manutenção de um equilíbrio reflexivo entre nossas performances particulares e regras indutivas só pode ocorrer na dimensão da nossa experiência.

Suponhamos por um momento que Goodman seja bem-sucedido em resolver a dificuldade inicial do antigo enigma com o recurso do equilíbrio reflexivo. Resta ainda o que ele chama de *novo* enigma. Para entender esse caso, vamos usar o exemplo do próprio Goodman: imagine que você está incumbido com a tarefa de descrever as esmeraldas que tira de uma urna até um tempo futuro *t*. Você pega a primeira esmeralda e nota, antes de *t*, que ela é verde. Pega a segunda e nota, antes de *t*, que ela também é verde. Imagine que a urna é grande e que, depois de um número suficientemente longo de esmeraldas verdes examinadas antes de *t*, você forma a crença de que *todas as esmeraldas são verdes*. O enunciado em itálico tem a forma de uma lei da natureza, digamos, trata-se de um enunciado *legiforme*. Goodman, então, introduz um novo predicado: *ser verzul*. Um objeto é verzul se ele é verde antes do tempo *t* e azul depois de *t*. O problema é que, a partir das suas observações, dado que todas elas ocorreram antes de *t*, você tem a exata mesma evidência para concluir que (1) todas as esmeraldas são verdes e que (2) todas as esmeraldas são verzuis. Todavia, essas hipóteses são fundamentalmente diferentes e oferecem previsões incompatíveis, visto que uma esmeralda depois de *t*, segundo (1), deverá ser verde enquanto, já segundo (2), deverá ser azul. O problema é que, ao passo que (1) é um enunciado legiforme, (2) não é — mas como seria possível diferenciar um enunciado legiforme de um enunciado não legiforme? Note que ambos enunciados têm a forma de uma quantificação universal ('todo *A* é *B*'), portanto, são indistinguíveis do ponto de vista estritamente lógico. O novo enigma da indução, então, diz respeito a quais enunciados são confirmados indutivamente em detrimento de outros, visto que alguns, como o enunciado acerca das esmeraldas verdes, são aceitáveis; enquanto

outros, como o enunciado acerca das esmeraldas verzuis, não. A resposta de Goodman ao novo enigma é simples, mas pode parecer insatisfatória: não há nenhuma razão além das nossas próprias práticas inferenciais para determinar que (1) é legiforme em detrimento de (2), pois de fato usamos predicados como *ser verde*, mas não usamos predicados como *ser verzul* — não há, por assim dizer, histórico do uso desse predicado. Assim como Hume havia tentado resolver o antigo enigma pela força do hábito, a solução de Goodman ao novo enigma também reside no fato de que habitualmente usamos alguns predicados e não outros, de modo que não há nenhuma razão ulterior que permita discriminar bons enunciados, que são confirmados indutivamente, de maus enunciados, que não deveriam ser confirmados.

Subdeterminação e abdução

Imagine que você é uma profissional da área de saúde e que está examinando um paciente que apresenta os sintomas típicos do novo coronavírus. Você quer saber se a seguinte hipótese (*H*) é verdadeira: *o paciente P tem coronavírus*. Imagine que você não dispõe de um teste confiável que permita descartar falsos positivos (isto é, casos em que a pessoa parece ter a doença, mas de fato não a tem). Você faz uma lista de observações dos sintomas apresentados, e a melhor explicação para essas observações é a verdade de *H*. Ou seja, você infere abdutivamente que *H* é verdadeira a partir das evidências coletadas. O problema é que a hipótese *H* pode ser *subdeterminada* pela evidência disponível. Ou seja, a evidência pode confirmar *H*, mas poderia confirmar também outra hipótese *H**, digamos, que *o paciente P tem um tipo ainda mais novo de doença viral* (eu honestamente espero que não!). É possível que *H** seja verdadeira e que seja uma explicação ainda melhor das suas observações prévias. Desse modo, as evidências não permitem determinar *H* como verdadeiro em detrimento de outras hipóteses rivais como *H**.

É claro, alguém poderia argumentar que não há razão nenhuma para pensar seriamente na verdade de *H**. Concedendo isso, podemos

formular outro argumento que ameaça a viabilidade de inferências abdutivas: qualquer inferência abdutiva envolve a seleção de uma explicação, idealmente a melhor, entre um conjunto de explicações possíveis. No entanto, para qualquer explicação escolhida, é possível que existam outras explicações que até o momento *não foram concebidas*, isto é, que não foram avaliadas até o ato de selecionar aquela que seria supostamente a melhor entre elas. Ou seja, pode ocorrer que uma inferência abdutiva apenas selecione a melhor explicação *entre as explicações consideradas*, que podem não ser todas. Assim sendo, é possível que haja explicações melhores que permanecem ainda inexploradas. Se esse for o caso, como saberemos se uma inferência abdutiva qualquer foi bem-sucedida? Em especial, como saberemos que as teorias científicas que resultam de inferências abdutivas estão corretas?

Esse argumento, com efeito, ganharia mais força se pudéssemos constatar que, historicamente, casos de subdeterminação desse tipo foram frequentes. Assim, teríamos boas razões para acreditar que as melhores teorias científicas atuais estariam igualmente ameaçadas, porque não passariam das melhores explicações *entre as concebidas pelos cientistas* dos fenômenos observando. Essa é uma possibilidade que investigaremos no final deste capítulo.

Ciência e aquilo que tenta passar-se por ciência

Talvez a tarefa mais urgente para a filosofia da ciência seja oferecer um critério que permita discriminar a prática legitimamente científica daquilo que tenta passar-se por ciência, a *pseudociência*. O tema é urgente porque a pseudociência está por todo lado, da indústria de homeopatia que movimenta quantias vultosas de dinheiro (e que leva pessoas desavisadas a evitarem tratamentos tradicionais e a buscarem tratamentos ineficazes) até figuras políticas que estão efetivamente condenando as próximas gerações a habitar um mundo desolado porque contrariam a opinião amplamente consensual dos climatologistas. Aquelas figuras argumentam — talvez estupidamen-

te, talvez sadicamente — que não devemos nos preocupar com as mudanças climáticas porque estiveram na Europa no verão e "até que fez frio" (pense consigo mesmo: você comeu algo hoje? Se sim, isso não significa que não existe fome no mundo, não é?).

Mas antes da pseudociência ter se tornado o principal motor da crise de racionalidade que presenciamos, o problema de delimitar a ciência foi bastante popular no século XX, sobretudo a partir dos avanços do empirismo (ou positivismo) lógico. Para os filósofos do Círculo de Viena, nós vimos anteriormente, compromissos metafísicos deveriam ser evitados porque enunciados metafísicos não fazem sentido, ao contrário de enunciados científicos. Sua preocupação, portanto, era oferecer um critério que permitisse discriminar ciência de metafísica (não exatamente uma pseudociência como o negacionismo climático, o movimento antivacina, o terraplanismo, etc.), e para isso usavam o verificacionismo. Nessa concepção, um enunciado cientificamente legítimo é *verificável pela experiência*. Parece plausível que enunciados típicos da ciência devem poder ser descobertos como verdadeiros, enquanto enunciados de atividades não-científicas, como a metafísica, não se prestem à verificação.

No entanto, o verificacionismo como um critério de demarcação da ciência tem um grande problema. Enunciados científicos frequentemente têm a forma de uma quantificação universal do tipo 'todo sódio em forma sólida em contato com água nas condições ideais gera uma combustão'. Mais abstratamente, enunciados desse tipo são da forma 'todo *A* é *B*', quantificações universais. Mas é obviamente impossível *verificar* um enunciado desse tipo, porque ele fala sobre *todos* os *A*s que eles são *B*s — não apenas dos *A*s observados até agora, mas de aqueles que existiram no passado remoto e que existirão no futuro distante. Não é viável observar *todo* pedaço de sódio em contato com água e, portanto, não é possível verificar que 'todo sódio em forma sólida em contato com a água nas condições ideais gera uma combustão'. Se esse for o caso, enunciados obviamente científicos como esse teriam de ser considerados *pseudocientíficos*, o que é obviamente absurdo.

Poderíamos tentar suavizar a necessidade de verificação pensando em *confirmação empírica*, de modo que a constatação de um *A* que é *B* confirmaria (mas não verificaria) a hipótese de que todo *A* é *B* — isso, é claro, pelo menos até a ocorrência de um *A* que não é *B*. É bastante plausível, com efeito, que a observação de um corvo preto, por exemplo, confirme (pelo menos provisoriamente) a hipótese de que todo corvo é preto, e que a observação de mais corvos dê maior suporte indutivo àquele enunciado. Um problema apresentado pelo filósofo alemão Carl Hempel é que um enunciado como (1) 'todo corvo é preto' é logicamente equivalente ao enunciado (2) 'tudo que não é preto não é corvo'. Isso pode não ficar óbvio em linguagem ordinária, mas posto em uma linguagem minimamente formalizada, nós temos o seguinte:

1. Para todo *x*, se *x* é corvo, *x* é preto.
2. Para todo *x*, se *x* não é preto, *x* não é corvo.

Os enunciados acima são o que chamamos de *contrapositivos*, porque eles têm as mesmas condições de verdade. Ou seja: somente se (1) é verdadeiro, (2) é verdadeiro. Se qualquer uma das duas frases for falsa, a outra também é. O problema é que podemos confirmar (2) com a observação de algo que não é corvo nem preto, como um gato branco. Se confirmamos (2) com a observação de um gato branco, então confirmamos (1) com essa mesma observação, porque (1) e (2) têm as mesmas condições de verdade. Esse problema, conhecido como *Paradoxo dos Corvos* em virtude do exemplo utilizado por Hempel, parece jogar a pá de cal em um critério confirmacionista de enunciados científicos.

Mas se enunciados científicos não podem ser verificados porque têm a forma de quantificações universais, eles ainda podem ser *falseados*. Isso significa que, para um enunciado do tipo 'todo *A* é *B*' é possível descobrir que ele é falso pela observação de um *A* que *não é B*. Você pode descobrir, por exemplo, que é falso que todo corvo é preto ao descobrir um corvo albino. Claro, esse é um exemplo sim-

plificado que beira a trivialidade, mas ele representa uma assimetria interessante com respeito aos procedimentos de verificar e de falsear enunciados universalmente quantificados. Isso não passou desapercebido pelo filósofo austríaco Karl Popper (aquele que estava na outra ponta do atiçador de lareira naquele evento com Wittgenstein), cuja filosofia é até hoje apreciada entre alguns círculos científicos.

O fundamento da filosofia da ciência popperiana é que a ciência não avança através da descoberta de verdades, mas através de conjecturas e refutações. Do ponto de vista lógico, segundo Popper, o pensamento científico *nunca* envolve inferências indutivas, o que lhe exime de resolver o problema da indução. Ou seja, ao invés de começar por observações — porque, afinal de contas, não faria sentido procurar por evidências sem saber qual hipótese se quer confirmar, isto é, sem saber *do que* são essas evidências —, a cientista começa pela concepção de uma hipótese como 'todo *A* é *B*'. A partir disso, ela segue à procura de uma refutação, nesse caso, um *A* que não é *B*. Se tal observação for feita, a hipótese inicial terá sido refutada e deverá ser substituída por uma hipótese mais apta, uma que dê conta das observações anteriores. Enquanto não é encontrado um caso que refute a hipótese inicial, ela é *corroborada*, mas nada disso permite inferir que ela seja *verdadeira*. Com efeito, a tarefa da ciência não seria, na visão de Popper, de descrever o mundo verdadeiramente, mas de avançar ao descobrir como o mundo *não é*. A cada refutação, uma nova hipótese é concebida e, desse modo, afasta-se aos poucos da falsidade. Quanto mais informação tem um enunciado, mais observações podem falseá-lo — compare, por exemplo, os enunciados 'a Terra orbita o Sol' e 'a órbita da Terra ao redor do Sol é elíptica'. Não é qualquer observação da órbita terrestre que falsearia o primeiro enunciado, ela poderia, por exemplo, ser elíptica ou (bizarramente) quadrangular. Mas se observássemos que a órbita terrestre não é elíptica, o segundo enunciado seria falseado, diferentemente do primeiro. Por isso, o segundo enunciado é mais informativo e, portanto, preferível do ponto de vista do progresso científico.

Popper é às vezes injustamente associado aos positivistas lógicos por defender que a ciência tem um único método, a saber, a busca por observações que refutem as hipóteses com as quais cientistas trabalham. Segundo Popper, isso vale tanto para as ciências naturais, como a física, a química e a biologia, quanto para as ciências sociais e humanas. Como nós vimos, a tese de que há apenas uma metodologia científica é uma tese tipicamente positivista, mas o positivismo também depende da seguinte especificação: o método científico é *indutivo* (para os positivistas modernos) ou *verificacionista* (em especial para os positivistas lógicos do Círculo de Viena). Ou seja: para o positivismo clássico, a ciência parte de observações particulares à inferência indutiva de hipóteses gerais, enquanto para o positivismo lógico, a marca da ciência é que seus enunciados são verificáveis. A filosofia de Popper nasce justamente da negação do verificacionismo e, em virtualmente toda sua obra, Popper enfaticamente rejeita a utilidade, e às vezes até mesmo a existência, de inferências indutivas na prática científica. Portanto é profundamente descabido acusá-lo de ser um positivista. Naquele que talvez seja o pior debate famoso da filosofia — não por causa de Popper, mas pelos seus interlocutores — ocorrido em Tübingen, na Alemanha, em 1961, encontramos Popper cuidadosamente elencando 27 teses da maneira mais clara e acessível sobre as ciências humanas. Ele afirma, por exemplo, que, embora seja impossível para um pesquisador em ciências humanas livrar-se dos seus valores durante sua investigação científica, as humanidades ainda podem ser caracterizadas como objetivas em virtude da abertura à crítica aos seus resultados, assim como nas demais ciências. Todo esse cuidado para, na sequência, ser acusado de positivista e indiretamente de indutivista de modo tortuosamente barroco, apenas no limiar do compreensível, por Theodor Adorno. O "debate" foi publicado como livro em 1969 sob o título, *Der Positivismusstreit in der deutschen Soziologie*, em alemão: *A Disputa Positivista na Sociologia Alemã*, e traduzido para inglês em 1976.

Segundo Popper, enunciados científicos seriam, então, por excelência falseáveis. Se aplicarmos essa ideia como critério para

separar a ciência da pseudociência, veremos que enunciados pseudocientíficos, por oposição, não seriam falseáveis. Pense, por exemplo, nos enunciados da astrologia, como em: 'pessoas de Áries têm sorte no amor'. O que exatamente isso poderia querer dizer? Que tipo de observação poderia ser usada para falseá-lo? Pensando desse modo, o critério popperiano parece plausível. Note, no entanto, que muitos enunciados pseudocientíficos não apenas são falseáveis como *foram falseados*, isto é, são falsos. Por exemplo, 'um remédio se torna mais potente quanto mais ele é diluído' (o segundo princípio da homeopatia) e 'a ação humana não interfere nas mudanças climáticas' (o credo dos negacionistas climáticos) são dois enunciados comprovadamente falsos. Portanto, apesar de virtualmente todos os cientistas respeitáveis não considerarem essas afirmações legitimamente científicas, pelo critério de Popper, elas seriam científicas, visto que são falseáveis — tanto é que *foram* falseadas. Além disso, parece profundamente implausível que uma cientista não acredite na sua hipótese de trabalho, isto é, que não acredite que essa hipótese seja verdadeira, e que tente ativamente refutá-la — o que mostra que a visão de Popper, apesar de ser interessante do ponto de vista lógico, é inadequada do ponto de vista da descrição da prática científica. Popper está errado, mas sem dúvidas está errado com elegância.

Pouco depois do trabalho de Popper tornar-se o epicentro da filosofia da ciência pós-positivista, Thomas Kuhn desenvolveu uma perspectiva muito mais plausível, promissora e melhor informada cientificamente. Doutor em física e estudioso da história da física, Kuhn estava encarregado de ensinar história da ciência para as humanidades em Harvard nos anos 50 quando se deu conta de que o que Aristóteles chamara de *movimento* — a perda e o ganho de propriedades — não era o movimento da física contemporânea, de modo que a física aristotélica e a física contemporânea estão falando de coisas diferentes. Kuhn tinha, portanto, conhecimento de primeira mão da prática e da história da ciência ao propor, na sua grandiosa obra de 1962, *A Estrutura das Revoluções Científicas* (reimpressa com o acréscimo do posfácio em 1970), que a ciência não avança linearmente pelo

acúmulo de conhecimento. Pelo contrário, o desenvolvimento científico dá-se em períodos muito diferentes entre si, períodos de normalidade, de crise e de revolução.

Kuhn sugere que há períodos de *ciência normal* em que o progresso científico de fato dá-se pelo avanço linear e cumulativo, especificamente, pela solução de problemas que surgem durante a prática de uma disciplina científica. Nesse período, é como se cientistas resolvessem *quebra-cabeças* pela replicação de procedimentos que se mostraram exemplarmente bem-sucedidos no passado. Esses procedimentos exemplares de resolução de problemas, que são consensualmente reconhecidos pela comunidade científica como tais, são chamados por Kuhn de *paradigmas*. A análise aristotélica do movimento é um paradigma da física pré-moderna, a matematização dos campos eletromagnéticos por Maxwell é um paradigma da física moderna, e o geocentrismo ptolemaico é um paradigma da astronomia pré-copernicana, para citar alguns exemplos. Toda ciência normal ocorreria, portanto, sob a orientação de um único paradigma, que cientistas não buscam refutar nem verificar.

Com efeito, pode acontecer que o surgimento de um problema durante a ciência normal resista uma resolução tradicional, mas em um caso assim a primeira suspeita incidiria sobre o trabalho da cientista, não sobre a viabilidade do paradigma que orienta a sua pesquisa. Ou seja, de acordo com Kuhn (e contra Popper), a descoberta de uma observação que não se coaduna com as expectativas teóricas não leva à refutação do enunciado científico que descreve o paradigma. Pelo contrário, a comunidade científica busca maneiras de explicar essa anomalia *conservando* o paradigma. Contudo, casos desse tipo podem acumular-se, de modo a gerar uma *crise científica* no interior de uma disciplina. Às vezes ocorre que uma crise é resolvida sem grandes mudanças no paradigma, mas nem sempre isso é possível. Ptolomeu, por exemplo, observou que a trajetória dos planetas apresentava um breve recuo antes de avançar novamente. Da sua perspectiva geocêntrica, a explicação desse fenômeno envolveu postular *epiciclos*, isto é, pequenos círculos que caracterizam a rota-

ção dos planetas, que por sua vez fariam seu movimento ao redor de um círculo mais amplo, chamado de deferente. Com o progresso da astronomia ptolemaica, essa explicação tornou-se progressivamente mais complicada em virtude da observação de novos fenômenos, de modo a gerar uma intratabilidade teórica. Consequentemente, a confiança na viabilidade do paradigma ptolemaico tornou-se ameaçada, caracterizando o que Kuhn chamou aptamente de *crise científica*.

Em um momento de crise como esse, é comum o surgimento de um paradigma rival que pretende explicar ou evitar os fenômenos problemáticos observados até então — como, no caso da astronomia, foi o modelo heliocêntrico de Copérnico. Para o heliocentrismo, a necessidade de explicar o aparente recuo e avanço dos planetas é simplesmente dissolvida, porque, nesse novo modelo, a Terra também está em movimento ao redor do Sol, assim como os demais planetas. A revolução copernicana na astronomia, portanto, consistiu no surgimento de um novo paradigma que eventualmente substituiu o antigo ao resolver problemas que haviam se tornado praticamente intratáveis da perspectiva geocêntrica. A crise, na medida em que gera a substituição do antigo paradigma por um novo, caracteriza o que Kuhn chama de *revolução*. A revolução é um ponto de ruptura radical a partir do qual tem-se novamente um período de ciência normal, de conhecimento cumulativo. No caso específico da astronomia, cientistas passaram a orientar-se não mais pela visão ptolemaica, mas pela visão copernicana na solução de quebra-cabeças empíricos, como a previsão da órbita de corpos celestes a partir de um ponto de observação móvel.

Se aplicarmos as ideias centrais de Kuhn para separar ciência de pseudociência, podemos argumentar que a ciência (em período normal) é orientada por um paradigma consensualmente aceito pela comunidade científica, enquanto a pseudociência apoia-se em ideias alternativas (como antigos paradigmas, ou propostas que nunca foram avaliadas como possíveis paradigmas). Por exemplo, *hoje* não restam dúvidas de que a química é uma ciência bem estabelecida, enquanto a alquimia é vista por cientistas como uma pseudociência. Pode ser

o caso que, em algum momento no passado, a alquimia tenha sido considerada uma empreitada legitimamente científica. De modo semelhante, especulações vazias como as que constituem a chamada "teoria" do design inteligente, o negacionismo climático, o terraplanismo e o movimento antivacina estão fora da prática científica normal e, portanto, são obviamente pseudocientíficas. A ideia kuhniana é promissora, porque ela desloca a questão sobre se *enunciados* são pseudocientíficos para a questão sobre o que a *comunidade científica* faz, e a resposta a essa pergunta, portanto, passa a considerar como a ciência é de fato feita em dado tempo.

Há, no entanto, alguns problemas com a perspectiva historicamente orientada desenvolvida por Kuhn. Um deles é a suposição de que pode haver apenas um paradigma vigente por disciplina científica durante do período de ciência normal, uma impressão que pode ser gerada pelos exemplos por ele elencados. Mas observamos hoje que, dentro de uma mesma disciplina científica, há vários paradigmas competindo concomitantemente. Esse problema, no entanto, pode ser resolvido com alguns ajustes à posição original de Kuhn. Além disso, a ideia de que o que constitui a ciência é aquilo que os cientistas fazem (durante a ciência normal) é um caso flagrante de circularidade explanatória, e defender esse critério de demarcação científica exigiria explicar por que ela se trataria de uma circularidade não-viciosa.

Por fim, segundo Kuhn, em casos de crise, os proponentes de paradigmas rivais estão, em última análise, em uma disputa sobre qual é o modo correto de fazer ciência. Apesar de interessante, essa proposta tem um elemento profundamente controverso, pois os proponentes de diferentes paradigmas estariam usando *linguagens* diferentes, *vendo mundos diferentes*, de modo que as suas práticas científicas seriam, ao fim e ao cabo, o que Kuhn chama de *incomensuráveis*. Isto é, as virtudes teóricas e empíricas de um paradigma seriam incomparáveis com as virtudes teóricas e empíricas do paradigma rival, não haveria um critério objetivo que permitisse comparar qual dos dois é o melhor. Se esse for o caso, uma conciliação *racional* entre paradig-

mas divergentes seria impossível, pois não haveria ponto de acordo comum para ambos os lados, restando apenas o que Kuhn assemelha a uma *conversão religiosa.* A conversão científica ocorreria, portanto, a partir de elementos extra-epistemológicos, como a retórica dos cientistas, a elegância e a simplicidade de um candidato a paradigma, ou simplesmente o fato de que a velha-guarda científica tende a sair de cena com o tempo. O problema dessa perspectiva é que ela soa perigosamente relativista, pois, no que concerne aos conteúdos científicos, ambos paradigmas seriam "igualmente bons", e só poderíamos discriminá-los por fatores externos à ciência.

Seria possível salvar a teoria de Kuhn do relativismo? Uma via interessante consiste em rejeitar a tese da incomensurabilidade, o que por sua vez exige apontar que há características incontroversas fundamentais que permitem comparar resultados científicos relativos a paradigmas rivais, como o sucesso efetivo na resolução de problemas empíricos. Esse é um fator ao mesmo tempo epistemológico e pragmático, e não haveria razão para não o levar em conta na avaliação de qual modelo de fazer ciência é mais apto. Além disso, talvez seja interessante pensar nos critérios que pareciam ser extra-epistemológicos (elegância, simplicidade, etc.) como critérios legitimamente científicos, que qualquer um pode reconhecer dentro de certos parâmetros. Assim sendo, cientistas que trabalham em paradigmas diferentes não estariam vendo mundos diferentes, trabalhando com linguagens radicalmente incompatíveis, pois haveria um ponto de contato que permitisse comparar seus resultados.

A dificuldade de demarcar ciência da pseudociência parece requerer que tenhamos uma boa definição, ou ao menos um bom esclarecimento, do que é *o* método científico. Com efeito, o indutivismo, o verificacionismo, o falsificacionismo e a resolução de quebra-cabeças sob o plano de fundo de um paradigma são tentativas de lançar luz sobre o método característico das ciências. Mas talvez seja melhor abandonar a suposição de que existe *um único* método científico, um método que é posto em prática por todas as ciências. Talvez ciências particulares empreguem métodos diferentes — talvez o modo como

um biólogo faz previsões sobre uma população seja diferente do modo como um físico testa uma teoria e do modo como um químico explica uma reação. Esse seria um caso de pluralismo metodológico. Agora, se o pluralismo metodológico estiver correto, deveríamos dizer que não é possível separar ciência de pseudociência — porque, supostamente, *vale tudo*? É óbvio que essa é uma conclusão que queremos rejeitar, então é importante procurar por algum critério de demarcação que não descarrilhe em um relativismo grotesco. Talvez o que separe a ciência genuína da farsa pseudocientífica sejam características muito gerais que não necessariamente dizem respeito ao método, mas à postura dos cientistas em oposição à sua contraparte embusteira. Por exemplo: cientistas devem ser sensíveis a novas evidências, devem seguir procedimentos amplamente aceitos como adequados, devem exibir a disposição crítica para reavaliar teorias de acordo com o avanço empírico, devem reconhecer a autoridade epistemológica de especialistas — e, além disso, a boa ciência não pode prescindir de procedimentos de avaliação institucionalizados, como a revisão por pares e a replicabilidade de experimentos científicos. Essas não são características que encontramos em grupos anticientíficos.

Ciência gera mesmo conhecimento?

O trabalho visionário de Kuhn deu início a uma nova perspectiva na filosofia da ciência, pois voltou a atenção dos filósofos à história e à sociologia da ciência. Mas esse novo olhar também levantou uma suspeita nova: será mesmo que as nossas melhores teorias científicas são verdadeiras? Razões históricas parecem mostrar que não são. Deveríamos dizer, então, que a ciência não produz conhecimento?

O argumento clássico em defesa da perspectiva de que as nossas melhores teorias científicas são verdadeiras, a posição conhecida como *realismo científico*, é chamado de *Argumento do Milagre*. Esse argumento afirma que a melhor explicação para o sucesso das nossas teorias científicas é o fato de que elas são verdadeiras, ou ao

menos aproximadamente verdadeiras — no sentido de que, ainda que existam explicações e enunciados falsos no quadro geral de uma teoria bem-sucedida, a teoria ainda é majoritariamente composta por enunciados verdadeiros. O sucesso em questão diz respeito, por exemplo, aos fatos de que as teorias aceitas presentemente explicam de maneira suficientemente robusta os fenômenos observados, de que fazem previsões acuradas e de que resistem aos testes com o passar do tempo. Dadas essas virtudes, como o nome do argumento sugere, seria um milagre se elas não fossem ao menos aproximadamente verdadeiras.

Note, no entanto, que as razões históricas mencionadas acima mostram que, sempre que uma teoria científica foi amplamente aceita pela comunidade especializada da sua época — por exemplo, a física aristotélica, a astronomia ptolemaica, a teoria do flogístico e a teoria da geração espontânea —, essa teoria também tinha sucesso empírico e também foi eventualmente substituída. Se, no entanto, mesmo as melhores teorias do passado foram eventualmente rejeitadas, não obstante o seu sucesso empírico, é razoável inferir indutivamente que as melhores teorias atuais serão rejeitadas no futuro. Desse modo, não é seguro afirmar a verdade de uma teoria a partir do seu sucesso empírico. O que é pior: essa indução, que recebe o nome adequado de *indução pessimista*, sugere que as nossas teorias atuais são provavelmente *falsas*. Isso motiva, portanto, um *antirrealismo científico*. E, além disso, se as teorias atuais são provavelmente falsas, então elas provavelmente não podem constituir conhecimento, uma vez que conhecimento implica sucesso (especificamente nesse caso, verdade), como mencionamos no começo do capítulo.

Esse é um impasse teórico de difícil resolução. Por parte dos realistas, uma tentativa de resposta à indução pessimista consistiu em mostrar que, em todas as teorias empiricamente bem-sucedidas, alguns elementos comuns permaneceram através da substituição por teorias melhores, e que esses elementos seriam justamente aqueles responsáveis pelo sucesso empírico tanto das antigas teorias científicas quanto das novas. Por exemplo, a antiga teoria do flogístico foi

proposta inicialmente no século XVII como uma explicação da combustão. Segundo essa teoria, todos os materiais inflamáveis contêm um elemento chamado de *flogístico*, que é liberado durante a combustão. Assim, em ambientes fechados (como dentro de uma redoma), a combustão de um material cessaria porque o ambiente ficaria supersaturado de flogístico. Lavoisier, no século XVIII, mostrou que o que explicava as reações observadas anteriormente é a presença do oxigênio, um elemento que havia sido recentemente descoberto. Portanto, argumentam os realistas, o que havia de bem-sucedido na teoria do flogístico é o fato de que o nome 'flogístico' na verdade refere-se ao que conhecemos hoje por oxigênio.

Essa resposta, portanto, consiste em reconhecer que existem elementos preservados através da substituição teórica, como alguns termos técnicos que os cientistas usam nas suas explicações. O termo 'flogístico', no exemplo acima, na verdade refere-se ao oxigênio, embora seja usado em enunciados falsos, como 'todos os materiais inflamáveis contêm flogístico'. É claro, a teoria não estava completamente correta, porque o oxigênio não é liberado durante a combustão, nem está contido nos materiais inflamáveis, mas está no ar — razão pela qual conter um material em combustão por uma redoma impede a circulação de oxigênio, que é condição necessária para a combustão. Desse modo, com o surgimento de uma teoria melhor, parte da teoria é preservada (o que explica o seu sucesso empírico), parte é abandonada — e assim o realismo sairia reivindicado. Apesar de promissora, contudo, essa resposta traz consigo um problema ainda maior. Vamos supor, como é razoável, que os cientistas sejam os especialistas nos assuntos com que estão tratando. Ao deparar-se com um problema, uma cientista propõe uma teoria composta de vários enunciados, cada um desses enunciados contendo alguns termos técnicos que se referem a certas entidades para que a teoria faça sentido. Mas, se o realista estiver certo na sua resposta acima, a cientista, que é a especialista em questão, *não sabe a que se referem os termos que ela própria usa*! Pois, afinal de contas, eventualmente sua teoria será rejeitada (ao menos parcialmente), e descobrir-se-á que os termos que ela

empregava referem-se a outras entidades. Parece absurdo que uma especialista num assunto não seja a pessoa na melhor posição para dizer sobre o que se trata o assunto em questão. E, no entanto, é isso que o realista está sugerindo.

Fortuitamente, há alternativas em defesa do realismo científico. Uma delas consiste em negar que há tantos casos de substituição teórica quanto a indução pessimista supõe. Claro, *algumas* teorias que foram amplamente aceitas no passado eram bem-sucedidas do ponto de vista empírico, e essas teorias, ainda assim, foram superadas por teorias melhores. Mas isso não quer dizer que *todas* teorias que foram superadas tinham o mesmo grau de sucesso empírico que suas sucessoras. Além disso, um realista poderia também negar que os procedimentos científicos do passado e de agora sejam *os mesmos* e que confiram *o mesmo grau* de confiabilidade às teorias atuais e passadas. Com efeito, a ciência é muito mais sofisticada hoje do que era há cem ou duzentos anos, e talvez a tendência é que existam menos casos de substituição teórica no futuro, e mais casos de refinamentos pontuais.

Mesmo que essas respostas realistas não encontrem sucesso, talvez possamos conceder que, por um lado, é mesmo possível que nossas teorias atuais sejam eventualmente rejeitadas e superadas por teorias melhores, mas que, por outro lado, isso não interfere no fato de que boas teorias oferecem *entendimento* sobre o mundo. De acordo com essa estratégia, poderíamos argumentar, o principal objetivo da ciência não é *descrever* o mundo verdadeiramente, mas *promover o entendimento* sobre como as coisas funcionam e permitir a intervenção bem-sucedida nos fenômenos observados. É claro, o entendimento científico pode ser refinado e ampliado com o progresso, mas isso não quer dizer que uma teoria que foi até então bem-sucedida seja um dia descoberta como *falsa*. Quer dizer apenas que ela promove um grau de entendimento da realidade inferior ao promovido pela sua sucessora. Assim, a indução pessimista estaria correta (de certo modo), pois temos boas razões para acreditar que o progresso da ciência nos levará a revisar a aceitação de certas teorias, mas o

argumento do milagre também estaria correto (de certo modo), pois as teorias funcionam porque elas nos permitem entender, intervir, prever, etc. a realidade, e isso não é um milagre.

Considerações finais

Nós começamos esse capítulo com observações muito gerais sobre o papel crucial que a ciência deveria ter nas nossas decisões políticas e também com alguns apontamentos sobre o que parecem ser sintomas de uma crise de racionalidade, como o surgimento de grupos anticientíficos, a disseminação de especulações estúpidas e perigosas, a falha sistêmica em reconhecer a autoridade epistêmica de especialistas e a resistência a explicações racionais. Apesar do cenário ser legitimamente desesperador, é a ocasião para que a filosofia desempenhe o importante papel de combater o cinismo sobre a racionalidade, o que nesse caso significa defender a ciência e o seu valor.

Leituras recomendadas

O livro clássico de Chalmers (1993) trata de muitos assuntos fundamentais de filosofia da ciência (como verificacionismo, indutivismo, falsificacionismo e revoluções científicas), mas é um livro antigo. O artigo de Papineau (2011) também cobre muitos assuntos que apresentamos aqui. Sobre o problema da demarcação, veja Hansson (2018). Uma compreensiva discussão acerca do problema da indução é encontrada em Castro e Fernandes (2014). O décimo capítulo de O'Brien (2013) também trata desse assunto, mas de modo bastante introdutório, bem como o sexto capítulo de Russell (2008). Sobre realismo e antirrealismo, veja Chakravartty (2017) e Castro (2014). Para uma claríssima apresentação sobre a história da filosofia da ciência no século XX, veja o livro de Moulines (2020).

6

O Bem, o Mal e o Feio

A filosofia, como eu argumentei no começo deste livro, é uma investigação contínua com o senso comum no sentido de que situações ordinárias e usuais podem desvelar perplexidades fundamentais, e responder a essas perplexidades requer um olhar propriamente filosófico. A ética talvez nos ofereça o mais incontestável exemplo de que questões absolutamente mundanas sobre o que as pessoas fazem têm como plano de fundo pressuposições filosóficas. Nós frequentemente nos deparamos com alegações da seguinte espécie: 'Fulano cometeu um erro, mas todo mundo faz a mesma coisa', ou: 'tudo que Beltrano fez deu errado, mas ele agiu com a melhor das intenções', ou ainda: 'Sicrano fez a coisa certa porque todos ficaram felizes'. Todas essas alegações indicam posicionamentos morais que cabem à filosofia investigar, como, por exemplo, se os critérios morais devem ser baseados em práticas culturais, ou se agir corretamente consiste em ter uma boa intenção, ou ainda se o que importa do ponto de vista moral são as consequências de uma ação.

Você pode estranhar o uso de 'ética' e 'moral' no parágrafo anterior como expressões intercambiáveis. Essa é uma escolha que eu faço deliberadamente e com boas razões, como veremos em breve, mas ela pode parecer controversa à primeira vista, sobretudo porque há uma tradição que consiste na separação dos dois termos. Essa separação às vezes é feita com base na teoria ética hegeliana, pois ela distingue o conceito de "eticidade" ou "vida ética" (em alemão, *sittlichkeit*) do conceito de moralidade (*moralität*) — mas traçar

a distinção entre ética e moral com base em um neologismo arcaico e profundamente idiossincrático é obviamente arriscado. Talvez uma melhor razão para separar ética e moral deva-se à suposição de que ética trata do que devemos fazer, enquanto a moralidade supostamente se refere ao modo como as pessoas agem, duas coisas que não necessariamente coincidem. Com efeito, há uma razão etimológica relativa ao radical latino da palavra 'moral' para tratar a moralidade como relativa aos costumes ou ao modo de agir. Mas essa razão é ruim, porque a palavra 'ética' originalmente possui o mesmo sentido, só que em grego. Portanto, não há boas razões etimológicas para usar as expressões 'ética' e 'moral' como possuindo significados diferentes. É, contudo, importante notar que a suspeita de que ética e moralidade possuam significados diferentes não é *inteiramente* descabida, visto que ela é baseada na observação de que a ética não é necessariamente equivalente ao modo como as pessoas agem, isto é, ao que é socialmente aceito. Esse é o nosso próximo assunto — note, no entanto, que essa suspeita não autoriza a separação terminológica entre 'ética' e 'moral', visto que temos uma palavra ainda mais apropriada para referimo-nos aos nossos costumes, a saber... 'costumes'. Portanto, quando quisermos separar ética (e moral) da mera conduta socialmente aceita, vamos usar 'costumes', ou 'práticas culturais' ou ainda 'convenções sociais' para referirmo-nos ao modo como as pessoas se comportam.

A ameaça relativista

Talvez nenhum posicionamento filosófico seja tão disseminado hoje quanto o relativismo. Essa é, no entanto, uma concepção muito antiga, provavelmente inaugurada pelo sofista grego Protágoras, cujo lema teria sido "o homem é a medida de todas as coisas", já no século V a.C.. O relativismo manifesta-se na ideia de que as pessoas possuem "as suas verdades", como se a verdade pudesse ser relativizada para cada um, como comentamos no primeiro capítulo. Esse, com efeito, seria um relativismo *epistêmico*: a negação de que há uma diferen-

ça entre aquilo em que as pessoas acreditam e o que é verdadeiro, de modo que cada pessoa teria o seu conjunto de "verdades subjetivas", cada um possivelmente conflitando com as "verdades subjetivas" de outras pessoas. Não por acaso, eu especulo, a proliferação de factoides é tão eficaz, pois muitas pessoas relativizam a verdade ao que querem acreditar, especialmente se o conteúdo de um factoide reforça seus preconceitos e suas crenças anteriores. Esse conteúdo, por mais absurdo e descolado da realidade que seja — como, por exemplo, a distribuição de material escolar com formato de genitália ou o uso de fetos abortados como adoçantes de refrigerantes — não é recebido criticamente por pessoas que negam que as suas crenças *possam ser falsas*, pois, para elas, "cada um tem a sua verdade". A busca pela verdade, nessa concepção, seria trivializada para o mero relato das próprias opiniões.

O que nos interessa neste capítulo, no entanto, não é o relativismo epistêmico, mas o relativismo *moral*. A versão individualista (ou subjetivista) do relativismo moral diz que cada pessoa tem um conjunto de critérios morais, isto é, critérios que estabelecem o que é a coisa certa a se fazer, e que pessoas diferentes possuem conjuntos potencialmente incompatíveis de critérios morais. Por afirmar que critérios morais são subjetivos, essa versão individualista do relativismo moral implica o *antirrealismo moral*, pois nega que a moralidade tenha uma dimensão real, objetiva. Com efeito, isso parece fazer algum sentido à primeira vista, pois o enunciado 'a porta é branca', possui critérios objetivos claros de verificação (a porta em questão ser de fato branca) — mas não é tão óbvio que um enunciado como 'é errado matar' possa ser verificado da mesma maneira. Em resposta, um realista moral poderia observar que a suspeita acerca da dificuldade de verificação de um enunciado moral está baseada em uma concepção correspondentista da verdade. Com efeito, é (aparentemente, ao menos) simples estabelecer uma correspondência entre 'a porta é branca' e o fato de que a porta é branca (se ela for), mas não é tão simples estabelecer o mesmo tipo de correspondência entre um enunciado moral qualquer e a realidade, visto que conceitos morais como 'cer-

to', 'errado', 'bom', 'mal' etc., são normativos, não descritivos. Ou seja, esses conceitos dizem respeito a sobre como as coisas *devem* ser, não sobre como elas de fato são. No entanto, prosseguiria o realista, há razões independentes para rejeitar o correspondentismo, como vimos no capítulo 4, e uma concepção deflacionária da verdade evita a dificuldade de estabelecer uma correspondência entre um enunciado moral e o mundo. Nessa concepção, dizer que matar é errado é verdadeiro se, e somente se, matar é errado.

A versão individualista do o relativismo moral (ou relativismo subjetivo), no entanto, não é tão popular quanto a variedade de relativismo que identifica critérios morais com práticas sociais compartilhadas por um povo em certo tempo. Segundo a perspectiva cultural do relativismo moral (ou simplesmente relativismo cultural), portanto, há fatores objetivos para critérios morais, a saber, o fato de que uma sociedade aceita ou endossa certas condutas. Ou seja, o relativismo cultural não rejeita a objetividade de critérios morais: esses critérios são objetivos, mas relativos às práticas de cada povo.

Para entender o que o relativista moral quer dizer com isso, pense, por exemplo, no seguinte enunciado moral (*M*): 'é errado matar'. Segundo o relativismo cultural, o enunciado *M* deve ser reinterpretado de acordo com a seguinte regra de conversão, que gera o enunciado relativizado (*MR*): 'é errado matar para a sociedade *S*', em que '*S*' refere-se à sociedade da qual o falante faz parte. Assim, quando eu alego *M*, o que eu estou efetivamente alegando é que é errado, *para a sociedade da qual eu faço parte*, matar. Ou seja, um enunciado moral faz referência implícita aos costumes da sociedade da qual o falante faz parte, o que explicita que valores éticos seriam nada mais do que convenções sociais.

Essa regra de conversão é essencial ao relativismo cultural — e é uma ideia bastante persuasiva. Como nós mencionamos no segundo capítulo, a descoberta de povos com costumes diferentes dos costumes europeus prontamente levantou a suspeita de que todo tipo de afirmação moral seria implicitamente relativa a fatores que haviam até então passados despercebidos. Um europeu que considera, por

exemplo, o infanticídio abjeto estaria apenas manifestando uma preferência da sua sociedade. Dito por um esquimó, o enunciado 'o infanticídio é errado' seria falso, dada a constatação empírica de que muitas famílias esquimós sacrificam seus bebês. Do mesmo modo, nós presentemente dizemos que a escravidão é abominável, o que o relativista explicaria que remete ao fato de que nossa sociedade *hoje* reconhece a dignidade humana como um valor a ser preservado independentemente da etnia. Mas até 1888, um membro da sociedade brasileira que dissesse 'a escravidão é abominável' estaria proferindo um enunciado falso, porque a escravidão era aceita pelos seus pares. Assim, se pensarmos na regra de conversão acima, um mesmo enunciado moral *M* dito por sujeitos diferentes, *A* e *B*, pode ser verdadeiro se dito por *A* e falso se dito por *B*, a depender da sociedade a qual cada um pertence.

O principal apelo dessa variedade de relativismo é que ela vem acompanhada de um ar de aceitação das diferenças entre sociedades radicalmente distintas, um ar de correção política. O relativista moral apresenta sua visão de mundo, portanto, como aberta e acolhedora, o que faria dela moralmente superior às alternativas que defendem a não-relativização de valores morais. Contudo, isso é um erro temerário. Considere o seguinte cenário: imagine que você é uma pessoa que pertence a uma sociedade que cultiva a integridade ou incolumidade física das pessoas como um valor moral inegociável. Nessa sociedade, portanto, não seria aceitável atentar contra a integridade física de uma pessoa. Imagine agora que você tem diante de si uma jovem imigrante que foge de uma sociedade que pratica a mutilação genital feminina, uma prática violenta que ainda ocorre em alguns países africanos e asiáticos. Imagine que a própria jovem foi vítima da mutilação genital, e que você está encarregado(a) de fazer com que ela seja aos poucos integrada na sua sociedade. Agora considere as seguintes palavras e pense bem se você as teria dito para essa pessoa: 'O que eles fizeram com você é errado *aqui*, mas foi certo *lá*, porque a sociedade da qual você fazia parte naquele momento permitia isso.' O exemplo pode soar dramático, mas ele faz um bom ponto,

porque qualquer pessoa com o mínimo bom senso moral é capaz de perceber quão absurdo seria justificar algo tão atroz, como a mutilação genital feminina, com base no fato de que ele é sancionado pelas práticas correntes de algum lugar em algum tempo. O problema é que o relativista prioriza o respeito às práticas de outras sociedades em detrimento do respeito às pessoas.

Além disso, outro problema do relativismo cultural é que ele trivializa a questão sobre o que é moralmente correto. Diante de qualquer ato ou ação *A*, acerca do qual queremos saber se é a coisa certa a se fazer, se é moralmente permissível ou condenável, etc., de acordo com o relativista, haveria uma solução simples e inequívoca para decidir sobre a valoração moral de *A*, a saber: a constatação sobre se a sociedade em que nos encontramos aceita ou rejeita *A*. Imagine se você quer saber se você deve ajudar um morador de rua inválido ou uma criança faminta, e que você só tem condições para ajudar uma dessas pessoas. Esse parece ser um dilema moral genuíno — mas não para o relativista, pois bastaria consultar o que a sociedade aceita como correto. Mas obviamente não parece sensato descobrir o que seria a coisa certa a se fazer com uma pesquisa de opinião.

Ademais, falar do que "a sociedade" aceita como o fundamento moral pode parecer adequado em casos muito gerais e abstratos, mas, quando tratamos de casos concretos, como no exemplo do parágrafo anterior, é mais difícil entender o que exatamente o relativista moral quer dizer com "*a* sociedade". Onde começa e onde termina uma sociedade? É nas suas fronteiras físicas? Uma sociedade é redutível às pessoas que a compõem? Ou seriam as instituições que a delimitam? Se a maioria das pessoas do meu bairro tem uma opinião — digamos, ajudar o morador inválido — e se a maioria das pessoas do bairro seguinte tem a opinião contrária, a sociedade tem convenções sociais incongruentes? Ou esses bairros constituem sociedades diferentes? E uma sociedade se estende no tempo? Plausivelmente sim, mas agora imagine que uma pesquisa realizada ontem constatou que maioria das pessoas acha melhor ajudar a pessoa inválida. Mas que uma pesquisa realizada hoje constatou que houve uma mudança de opinião

em massa, e que mais pessoas agora acham correto ajudar a criança faminta. Esse tipo de mudança de opinião pública indicaria que os critérios morais de uma sociedade mudaram — e isso significaria dizer que seriam duas sociedades diferentes? O ponto neste argumento é que o conceito de *sociedade* não é mais claro e incontroverso do que os conceitos morais que o relativista tenta explicar, ou seja, sua estratégia argumentativa gera mais problemas do que resolve.

O relativista frequentemente critica o absolutista moral como se este dissesse que os valores morais *da sua sociedade* são corretos e os de outras são errados. Mas não é isso que ele está dizendo. Ele diz apenas que os valores morais são objetivos (contra o relativismo subjetivo) e invariáveis ou absolutos (contra o relativismo cultural), independente de contingências como o tempo e o lugar em que as pessoas vivem. Mas o absolutista não diz que a sociedade da qual ele faz parte instancia esses valores. Inclusive o absolutista pode dizer que as práticas *da sua* cultura estão erradas, algo que, plausivelmente, o relativista não pode fazer. Isso ocorre porque a própria regra de conversão relativista gera uma dificuldade insuperável para a teoria: imagine que você *discorda* de algo que é amplamente praticado na sociedade da qual você faz parte. Imagine que se trata de algo simples, como furar a fila. Se você dissesse 'furar a fila é errado', o relativista interpretaria o seu enunciado como sendo, na verdade, a alegação 'furar a fila é errado para minha sociedade'. No entanto, no exemplo em que estamos imaginando, você justamente estaria criticando algo que as pessoas de fato fazem na sua sociedade — ou seja, o relativista interpretaria seu enunciado como outro enunciado, que é *falso,* mas isso absolutamente não é o que você quer dizer! É claro, esse é um exemplo tosco, mas imagine que o que está em jogo é a escravidão para um abolicionista brasileiro antes de 1888. O próprio fato de que a escravidão foi abolida em 1888 mostra que pelo menos algumas pessoas consideravam-na errada, não obstante o fato de que ela havia sido socialmente aceita até então. Com efeito, se o relativismo estiver correto, torna-se impossível criticar a própria cultura do ponto de vista moral. Isso mostra que, ao contrário de ser uma posição etica-

mente superior, porque seria acolhedora das diferenças culturais, é uma suposição terrivelmente autocentrada, visto que concebe como moralmente legítimo tudo aquilo que uma sociedade, de um ponto de vista abstrato, sanciona.

Do ponto de vista lógico, o problema fundamental do relativismo é o fato de que o principal argumento em seu favor trata-se de uma falácia escandalosa, pois da constatação de que há sociedades com costumes radicalmente diferentes, não se segue que os critérios morais são relativos a cada uma delas. Com efeito, como observamos rapidamente no segundo capítulo, essa constatação empírica é perfeitamente compatível com a hipótese de que há muitas práticas sociais *erradas*, isto é, moralmente condenáveis (inclusive as nossas). Além disso, talvez seja possível argumentar que, contra o que assume o relativista, na verdade há ampla concordância acerca dos valores morais fundamentais através de sociedades diferentes, havendo discordância apenas sobre alguns pontos específicos, ou sobre circunstâncias excepcionais que permitem a violação de valores que, em circunstâncias normais, seriam preservados. Argumentar nessa direção mostraria que o argumento com base empírica em favor do relativismo, além de falacioso, parte de uma falsidade.

Do ponto de vista político, pode parecer surpreendente constatar que o relativismo encontra amparo não em posições progressistas e liberais, mas no conservadorismo. O conservadorismo político afirma que decisões políticas devem ser tomadas com base na experiência da tradição, e não em idealizações sobre a capacidade ética e intelectual humana — como acreditariam as posições rivais —, de modo a preservar estruturas e dinâmicas sociais consagradas pelo tempo. Se o conservadorismo aspira a qualquer esclarecimento ético além de político, isto é, sobre o que é *certo* fazer, ele pode ser sintetizado como a afirmação de que *o modo como as coisas foram é o modo como as coisas devem ser.* Um conservador minimamente mais esclarecido talvez esteja disposto a afirmar que o modo como as coisas foram é *ao menos aproximadamente* o modo como elas devem ser, admitindo apenas doses homeopáticas de progresso na sua visão de

mundo. Em ambos os casos, o conservador está comprometido com a chamada *falácia naturalista*, a inferência de como as coisas *devem* ser a partir de uma descrição sobre como as coisas efetivamente são. Como vimos, conceitos morais são normativos e não podem, portanto, ser reduzidos a descrições da realidade. De qualquer modo, o conservadorismo pode facilmente ser aproximado ao relativismo por duas razões relacionadas: em primeiro lugar, por basear-se em um ceticismo sobre as capacidades intelectual e moral humana e, por conseguinte, por negar que seja possível *descobrir* normas morais objetivas para além da mera tradição. Em segundo lugar, por aceitar que a história perfeitamente contingente de um povo ou de uma nação determine o conteúdo da moral. Visto que, a princípio, uma sociedade pode ser organizada de diversas maneiras (pode, por exemplo, instanciar uma organização fascista, liberal ou comunista, só para citar alguns exemplos), de modo a manter-se estável e estabelecer uma tradição, então pode haver diversas regras morais incompatíveis entre si. Em resumo, a moral seria relativa à tradição, qualquer que ela fosse. É claro, esse não seria exatamente um relativismo dos departamentos de Antropologia, mas, digamos, um relativismo que toma o chá das cinco—um relativismo, não obstante. Se, no entanto, o conservador moral insistir que há critérios de organização social que permitem dizer que uma sociedade organizada de determinada maneira é moralmente preferível às alternativas (por exemplo, que uma organização liberal é moralmente preferível ao comunismo), então seu conservadorismo apela veladamente a critérios morais independentes da tradição, e a própria ideia conservadora é esvaziada de relevância teórica. Só sobraria o chá.

Até agora vimos que o relativismo moral é infundado e insustentável. Notemos, no entanto, que a rejeição do relativismo moral não nos deixa em uma situação muito mais confortável, porque agora temos de explicar como podemos descobrir ou constatar que algum critério moral é objetivamente adequado, seja ele socialmente aceito ou não. Esse é o nosso próximo assunto.

Ponderando as consequências

Não há dúvidas de que o experimento mental mais famoso da filosofia é o *Dilema do Bonde* (sem relação com funk carioca), inicialmente concebido pela filósofa britânica Phillippa Foot e mais tarde modificado por outra filósofa, a norte-americana Judith Jarvis Thompson. Não é exagero dizer que os trabalhos dessas duas filósofas foram cruciais para o retorno de questões éticas ao centro da tradição analítica na segunda metade do século XX. Foot foi uma das responsáveis por trazer a ética aristotélica de volta, e Thompson destacou-se em ética aplicada com o seu argumento em favor da permissibilidade do aborto, conhecido como *argumento do violinista*. A relevância dos trabalhos dessas autoras continua evidente, sobretudo pelo amplo alcance do Dilema do Bonde hoje em dia, transcendendo as barreiras acadêmicas, ocorrendo em séries de Netflix e até mesmo virando memes. Ainda assim, é curioso, mas talvez não surpreendente, que muitas vezes se refira a esse problema sem a menção a Foot e Thompson. De uma perspectiva otimista, no entanto, podemos pensar que esse dilema tornou-se tão basilar para a ética contemporânea que acabou sedimentado como um *dado*, tornando-se quase como um "fato filosófico" que dispensaria menção às autoras que primeiro o elaboraram.

Caso você ainda não conheça o Dilema do Bonde, a ideia é bastante simples. Imagine que você está observando os trilhos de um bonde em que há uma bifurcação logo adiante (este é o exemplo modificado por Thompson, no exemplo original de Foot, você seria o condutor). Você vê o bonde se aproximando e percebe que, se ele seguir o curso em que está, acabará por atropelar cinco trabalhadores que se encontram na pista (imagine também que eles são inocentes e que não há nada que eles possam fazer para evitar o acidente). Você nota que está ao alcance de uma alavanca capaz de mudar o caminho do bonde na bifurcação e, se fizer isso, o bonde acabará por matar apenas um trabalhador (também inocente e incapaz de evitar o acidente) que se encontra na outra pista. Ou seja: você pode deixar as coisas seguir como estão, e cinco morrerão, ou você pode intervir e causar a morte

de apenas um (salvando os outros cinco). Você tem o dever moral de puxar a alavanca?

Na discussão original, esse experimento mental serve para destacar nossas intuições morais acerca da (suposta) diferença entre matar e deixar morrer, mas aqui vamos focar no peso moral das consequências de uma ação. Se você constatou que você puxaria a alavanca, então você tem intuições *consequencialistas* em ética (lembre-se, estamos falando de intuições como disposições confiáveis para julgar ou avaliar algum evento que nos é apresentado). Há muitas variedades de consequencialismo, mas, de um modo geral, o consequencialismo é a tese de que o valor moral de uma ação reside unicamente nas suas consequências. Se, ao deparar-se com o dilema acima, você constatou que não puxaria a alavanca — que seria errado salvar cinco pessoas causando a morte de outra — então você tem inclinações *deontológicas* (do grego, *dever*). Assim como o consequencialismo, a deontologia admite variedades, mas aqui interessa saber que, segundo essa teoria moral, o valor de uma ação está na sua motivação, mais precisamente, na motivação em respeito ao dever. De acordo com isso, não seria uma motivação moralmente adequada decidir, intencionalmente, matar uma pessoa para salvar outras cinco.

Com efeito, essas duas posições (até então esboçadas de um modo muito geral) caracterizam a principal disputa ética da modernidade, que foi refinada apenas recentemente para admitir outros elementos e recuperar teorias pré-modernas. Ambos os lados da disputa moderna, os proponentes do consequencialismo e os proponentes da deontologia, supõem que todo fenômeno moralmente avaliável (isto é, que pode ser bom/correto ou mau/incorreto) é dividido em motivação, ação e consequências. Se, portanto, julgamos uma ação como moralmente correta, segundo o consequencialismo, isso se deve às consequências dessa ação. Segundo a deontologia, isso se deve à motivação que leva um agente a agir. Como mencionado acima, recentemente a ética contemporânea recuperou elementos das éticas pré-modernas, especialmente a antiga e a medieval, o que permitiu deslocar a avaliação moral de uma ação para a avaliação do *caráter* do agente. Uma

ação seria moralmente boa, por exemplo, se realizada por um agente virtuoso, ao passo que uma mesma ação, com as mesmas motivações e consequências, poderia ser considerada condenável se fosse realizada por um agente moralmente vicioso. De qualquer modo, aqui nós vamos focar na disputa entre consequencialismo e deontologia e, por conta disso, vamos nos ater ao modelo *motivação-ação-consequências*, de modo a colocarmos considerações sobre o caráter de agentes morais de lado. Note que há também teorias éticas que enfatizam o lado do paciente ou receptor de ações morais, não o lado do agente, mas nós não as avaliaremos aqui.

A principal teoria consequencialista, o utilitarismo, foi desenvolvida pelo filósofo britânico Jeremy Bentham (cuja cabeça mumificada é preservada até hoje no Colégio Universitário de Londres) no século XVIII, e é bastante persuasiva à primeira vista. A ideia fundamental do utilitarismo de Bentham (ou utilitarismo clássico, em oposição à versão mais refinada de John Stuart Mill), conhecida como princípio de utilidade, é que uma ação é moralmente correta se, e somente se, ele maximiza o bem. Na concepção utilitarista clássica, uma ação maximiza o bem se ela aumenta o prazer para todos e diminui a dor para todos. De acordo com isso, o bem é claramente entendido relativamente a *prazer* e *dor*, razão pela qual o utilitarismo clássico é uma teoria ética *hedonista*. Além disso — e talvez daí advenha seu maior apelo — o prazer e a dor de cada uma das pessoas afetadas por um ato são considerados indiscriminadamente, isto é, com igual peso, de modo que uma ação não se tornaria correta, por exemplo, se ela maximizasse apenas o bem-estar do agente, ou de um grupo de pessoas sob o preço do sofrimento das demais.

Pelo menos à primeira vista, o utilitarismo clássico se coaduna naturalmente com algumas de nossas intuições morais sobre casos ordinários e imaginários. Uma pessoa que atenta contra a vida de outra, por exemplo, para roubar a sua carteira, está causando muito mais sofrimento do que bem-estar para todos os envolvidos (nesse caso, na ausência de mais informações, podemos supor que o único beneficiário seria o autor do crime) — razão pela qual sua ação é

moralmente condenável segundo o utilitarista. De modo semelhante, parece-nos no mínimo ofensivo que os diretores da Vale do Rio Doce recebam um bônus relativo ao ano de 2019 de 19 milhões de reais, gerando extrema felicidade para poucas pessoas, justamente aquelas que seriam direta ou indiretamente responsáveis pelo desastre de Brumadinho, ocorrido naquele ano, enquanto as famílias dos 240 mortos nesse desastre ainda esperam compensação. É também intuitivo afirmar que os bilionários Bill e Melinda Gates agem corretamente ao doarem parte da sua imensa fortuna para pesquisas em saúde visando países em desenvolvimento, pois assim estão diminuindo muito mais sofrimento do que estariam se apenas estivessem acumulando mais fortuna para si. Um caso oposto é o de Jeff Bezos, o dono da Amazon. Bezos não apenas ficou mais rico durante a pandemia, mas o faz sem permitir que os funcionários da Amazon tenham condições de trabalho adequadas. Cerca de 70% desses funcionários relatam que se veem obrigados a urinar em garrafas de plástico porque a demanda é muito alta, e eles não têm tempo de deslocar-se até o banheiro nos enormes galpões da companhia. Intuitivamente, a conduta de Bezos é moralmente inadequada, o que o utilitarista explicaria através do fato de que o prazer que ele adquire pela economia em efetivamente proibir funcionários de ir ao banheiro é ínfimo, especialmente se comparado com a sua riqueza já acumulada, enquanto milhares de trabalhadores sofrem diariamente. No caso do bonde, se você decidiu puxar a alavanca para poupar os cinco trabalhadores, mesmo que isso resultasse na morte de uma pessoa, talvez você tenha pesado as consequências dessa ação, computando o bem-estar geral dos envolvidos. Se os cinco trabalhadores tivessem sido mortos, haveria mais sofrimento do que se apenas um trabalhador morresse — além das cinco vidas perdidas, cinco famílias teriam sofrido, e não apenas uma.

Não é surpreendente que o consequencialismo tenha sido a principal teoria moral dos empiristas modernos, pois, afinal de contas, as consequências de uma ação são empiricamente verificáveis — ao contrário daquilo que move o agente. Com efeito, a ideia de um cálcu-

lo objetivo de prazeres e dores pode parecer promissora, sobretudo como uma tentativa de afastar a ameaça relativista. Mas talvez as coisas não sejam tão fáceis assim: como, afinal de contas, poderíamos medir os prazeres causados por uma ação? E supondo que isso seja possível — supondo que tenhamos uma "escala" de prazeres e dores — quando as consequências de uma ação começam e quando elas terminam? Imagine que uma ação moral *A* maximizou a felicidade e minimizou a dor de todos os envolvidos em um primeiro momento. Essa ação, segundo o utilitarista, seria moralmente correta. Mas imagine agora que um desdobramento imprevisível fez com que a ação *A* causasse indiretamente mais dor do que se não tivesse sido realizada. Então *A* deveria ser considerada imoral? Talvez o utilitarista precise de uma restrição para as consequências imediatas da ação, de modo que consequências posteriores não seriam levadas em conta.

Se, contudo, o utilitarista aceita a restrição de consequências imediatas para a avaliação moral de uma ação, então sua teoria encontra ainda mais problemas. Um desses problemas decorre da melhor resposta utilitarista a outra dificuldade clássica. Muitos críticos perceberam que o princípio de utilidade sancionaria cenários moralmente absurdos: por exemplo, imagine que, para maximizar a felicidade dos habitantes de uma cidade, uma criança inocente é mantida em cárcere privado, sendo torturada diariamente. Conceda por um momento que exista uma (bizarra) relação causal entre a tortura dessa criança e a elevação ao máximo do prazer dos habitantes da cidade. O sofrimento dessa única criança pode ser tremendo, inimaginavelmente profundo — mas, mesmo assim, os prazeres provocados por esse ato abjeto em muito superariam a intensidade do sofrimento. O utilitarista parece ver-se forçado a aceitar que esse caso hediondo é moralmente correto, *contra as nossas intuições morais*. Há, no entanto, uma resposta a essa dificuldade clássica, a saber: *ao longo do tempo*, o conhecimento público de que há uma criança mantida em cárcere privado acabará gerando a suspeita de que outra pessoa eventualmente ocupará o lugar da criança, de modo que as pessoas passarão a viver com medo. Portanto, *a longo prazo*, a ação de manter uma

criança em cárcere privado, mesmo que tenha como consequência imediata a maximização do prazer dos cidadãos daquela cidade, passará a ser uma fonte de sofrimento e, portanto, não é uma ação moralmente correta. Se aceitarmos essa resposta, no entanto, tendo em vista que a cadeia causal de uma ação é infinita e amplamente imprevisível, a ideia de avaliar uma ação a partir das suas consequências perde qualquer precisão.

O filósofo norte-americano contemporâneo Robert Nozick — senhor das melhores sobrancelhas da filosofia — apresentou uma poderosa objeção ao hedonismo implicado pelo utilitarismo clássico. Seu argumento parte do experimento mental da *máquina de experiência*. Imagine uma máquina superpoderosa capaz de gerar uma realidade virtual que confira aos seus usuários uma experiência de máximo prazer e de máxima realização sem nenhum sofrimento. A pessoa que voluntariamente decidisse ligar-se à máquina de experiência estaria agindo do modo moralmente correto segundo o utilitarismo clássico, *mesmo que na vida real* essa pessoa não tivesse nenhuma conquista pessoal, amizade genuína, entendimento sobre o mundo, enfim, nada que intuitivamente avaliamos com o bom e desejável. Esse resultado, no entanto, é profundamente contraintuitivo, o que gera mais um problema ao utilitarismo clássico na medida em que se compromete com o hedonismo. É possível, contudo, dissociar o utilitarismo do hedonismo, de modo a evitar esse tipo de problema. O próprio Nozick desenvolveu uma teoria que às vezes é chamada de "utilitarismo de direitos", segundo a qual uma ação é moralmente correta na medida em que maximiza, não o prazer, mas a preservação de certos direitos naturais (isto é, direitos logicamente anteriores a qualquer contrato social), como o direito à vida e à propriedade privada.

Sem qualificações substanciais (e talvez até mesmo com elas), o utilitarismo enfrenta a dificuldade fatal de sancionar ações que diríamos serem moralmente condenáveis, como a tortura e o sacrifício, conquanto essas ações sejam promovidas pelo bem da maioria. Essa crítica contundente ataca o cerne da teoria, não obstante a sua

plausibilidade inicial, pois mostra que o utilitarismo não faz justiça às nossas intuições morais. Uma alternativa é simplesmente rejeitar que nossas intuições morais tenham qualquer relevância para decidir em favor de (ou contra) uma teoria ética, de modo que seria um mero preconceito da nossa parte pensar que um sacrificiozinho aqui ou ali para o bem de todos seria realmente um problema. No entanto, nossas intuições são muito arraigadas e não conseguimos superá-las facilmente. Pense no profundo ultraje que muitos sentiram quando foi sugerido por algumas pessoas, em meio à pandemia do novo coronavírus — enquanto corpos eram amontoados em containers ou enterrados em covas coletivas —, que o comércio voltasse às suas atividades normais, mesmo que isso causasse diretamente algumas centenas de milhares de mortes. O gélido cálculo utilitarista consiste em ver a recuperação, ou melhor, a manutenção da economia como um bem cujo peso ultrapassa o do sofrimento de muitas famílias com mortes plenamente evitáveis. Insiste o utilitarista, com uma visão de mundo que é surpreendente compartilhada por aqueles que se veem como defensores de valores familiares: "os hambúrgueres de uma famosa rede de *fast food* não vão se vender sozinhos, então é melhor condenar os seus avôs à morte".

Motivação e deveres morais

Se pensar a moralidade a partir das consequências é problemático, então é plausível pensá-la do outro lado do modelo *motivação-ação-consequências*. De acordo com essa perspectiva, o que faz com que uma ação seja moralmente correta é a sua *motivação*, aquilo que leva um agente a agir. Com efeito, há no nosso vocabulário moral uma expressão que captura isso, a ideia de agir de acordo com o *dever*, de fazer o que *deve ser feito* — não obstante as consequências do que é feito. De fato, podemos pensar em muitos casos em que o que avaliamos é a motivação do agente, não o modo como as coisas se desenvolveram por consequência da sua ação. Por exemplo: você certamente está familiarizado com pessoas ou empresas que praticam atos de

caridade como forma de promoção em redes sociais. Elas podem muito bem estar efetivamente ajudando outras pessoas — mas, se a sua motivação para agir desse modo for, por exemplo, a manutenção de uma imagem pública favorável, parece intuitivo que elas não estariam agindo corretamente. Semelhantemente, podemos pensar que um estudante que entrega os coleguinhas colando em prova agiu corretamente, porque fez o que deveria ser feito, ainda que com isso tenha desagradado algumas pessoas e que inclusive perca alguns amigos. Em ambos os tipos de caso, o que estamos levando em conta é o dever moral. Como mencionamos, as teorias éticas baseadas nos nossos deveres morais são chamadas de *deontológicas*, e seguramente a mais influente delas é a teoria apresentada por Kant, principalmente na sua *Fundamentação da Metafísica dos Costumes*, publicada em 1785.

Apesar da redação no mínimo barroca, a *Fundamentação* é uma tentativa primorosa de mostrar a possibilidade do conhecimento objetivo e universal dos nossos deveres morais. Com efeito, se a moral aspira à universalidade — isto é, se os nossos critérios morais servem para todas as pessoas, em todos os tempos, independente de peculiaridades de cada um — então a moralidade não pode ser identificada a eventos empíricos, como as consequências das nossas ações. A proposta kantiana, portanto, é mostrar que os nossos deveres morais são universais porque são cognoscíveis *a priori*, independente de quaisquer peculiaridades empíricas como preferências individuais, e que acessamos os nosso deveres morais através do exercício da nossa faculdade racional. Daí a ideia, na terminologia kantiana, de que os nossos deveres morais são baseados em ou constituídos de uma *razão pura prática*. Para chegar a esse resultado, Kant cunha alguns termos técnicos e traça algumas distinções que devemos examinar com cuidado.

Em primeiro lugar, Kant distingue um agir legitimamente autônomo de um agir "heterônomo" — isto é, agir segundo motivações externas ao agente. Uma pessoa que age porque decide racionalmente agir de determinada maneira exibiria autonomia, enquanto uma

pessoa que é compelida a agir pelas suas preferências está sob um controle como que externo a si. Um agente que é movido pelo vício, por exemplo, é um caso claro de heteronomia, pois é como se ele seguisse ordens que não foram impostas por si, mas pelos seus desejos. A autonomia, ao contrário, consiste em agir de acordo com ordens *autoimpostas* pela própria razão. Kant, como bom iluminista, entende a razão como universal, igual para todos, de modo que o dever moral, por ser fundado na racionalidade, é igual para todas as pessoas que estiverem na mesma situação. Para nós, que somos seres cujas motivações para agir podem não ser apenas racionais, mas frutos dos nossos desejos e hábitos (o que Kant chama de *inclinações*), deveres morais apresentam-se como *limitações*, como restrições que nos impedem de agir visando apenas a nossa satisfação. Kant especula que um ente exclusivamente racional, como um anjo ou Deus, não perceberia a moral da mesma maneira — porque, por ser exclusivamente racional, esse ente não seria movido nem parcialmente pela satisfação de inclinações.

Kant sugere, portanto, que uma ação só pode ser moral se for autônoma no sentido forte, isto é, fruto de uma decisão racional. Com isso a ética kantiana marca uma distinção importante com respeito ao conceito de *liberdade*. Há, com efeito, ao menos dois conceitos bastante diferentes de liberdade. Às vezes dizemos de uma pessoa que ela é livre porque nada a impede de agir de determinada maneira. Um viciado, por exemplo, na medida em que não há ninguém que o impeça de sucumbir ao vício, pode ser considerado livre nesse sentido. Por tratar-se da ausência de restrições, essa concepção de liberdade é considerada a liberdade *negativa* e ela marca toda ética empirista, influenciando fortemente teorias políticas de inclinação liberal. A teoria moral kantiana, por outro lado, alega que há algo de profundamente contraintuitivo em considerar o indivíduo que sucumbe ao vício como *livre*, mesmo que ele seja desimpedido. O que há de errado no caso daquela pessoa é que suas ações não partem de uma escolha feita por si, elas não são fruto da sua *autonomia* em

um sentido forte. Daí a ideia de uma liberdade *positiva*: ser livre é, na concepção kantiana, ser autodeterminado.

Como então uma decisão racional pode ser moral? Kant notou que, na tomada de uma decisão, há dois usos da racionalidade: um uso *instrumental* e outro que podemos chamar de *constitutivo,* na falta de um nome melhor. Quando usamos a racionalidade instrumentalmente, tomamos uma decisão que não é necessariamente moral, por exemplo: se eu quero satisfazer meu desejo de comer uma pizza, eu raciocino que devo comprar uma pizza. "Devo comprar uma pizza" apresenta-se a mim como uma ordem que está sendo aqui condicionada ao meu desejo de comer pizza, razão pela qual Kant chama esse tipo de ordem de *imperativo hipotético*: '*Se* eu quero satisfazer meu desejo de comer pizza, eu devo comprar uma pizza'. Imperativos hipotéticos seriam, então, os procedimentos típicos para decisões de meios e fins, isto é, decisões acerca de quais são os melhores meios para atingir certos fins — como são os cálculos utilitaristas, em que o fim é a maximização do bem-estar. Mas esse tipo de ordem ou de imperativo não decide se a finalidade *por si só* é moralmente boa, pois, como vimos, um cálculo utilitarista é compatível com atos violentos, como tortura e sacrifícios. Kant usa outro termo para referir-se a uma ordem que determina, por si só, o conteúdo do dever. Trata-se do famoso *imperativo categórico*. Como o nome sugere, um imperativo categórico é uma ordem não condicionada, porque ela expressa o que deve ser feito — ponto. Por essa razão, o uso da racionalidade *constitui* ou até mesmo *constrói* o dever moral, ele não é condicionado a fatores empíricos e contingentes como gostos e preferências.

Note, no entanto, que uma pessoa pode agir *de acordo com o dever* sem agir *em razão do dever*. Se pensarmos novamente no exemplo das celebridades de Instagram que fazem caridade para manter uma imagem pública favorável, fica claro que elas agem de acordo com o que parece, ao menos intuitivamente, ser um dever moral — digamos, o dever de ajudar os menos favorecidos. Mas elas não fazem isso *em razão do dever*. Com efeito, o raciocínio que leva essas celebridades a fazer caridade é um imperativo hipotético: 'se eu quero

me manter bem visto(a) aos olhos da sociedade, devo fazer caridade'. Não parece ser o caso, no entanto, que devemos condená-las por agir *imoralmente*. O que ocorre é que suas ações não são moralmente *boas*, mas isso não quer dizer que sejam *erradas*. Para que uma ação seja moralmente boa, então, é preciso que ela tenha sido realizada em razão do dever. De modo semelhante, uma pessoa que age em razão do dever pode não tirar nenhum proveito disso — pode inclusive tornar-se mais infeliz processo. Mas é óbvio que não é necessário que ações moralmente positivas gerem apenas consequências desagradáveis. No melhor dos casos, uma ação moralmente correta terá, além da motivação com base no dever, consequências positivas.

Mas o que, afinal de contas, é o dever? Kant tenta mostrar o que devemos fazer através de três formulações específicas do imperativo categórico. Como essas formulações servem para todas as nossas ações, em todas as circunstâncias, há um alto grau de generalidade e de abstração envolvido — que aumenta a cada nova formulação — e uma boa maneira de entender esse procedimento é aplicando-o a casos concretos.

A primeira expressão do imperativo categórico é a *Fórmula da Lei Universal da Natureza*, seja ela: *aja como se a máxima da sua ação devesse tornar-se, através da sua vontade, uma lei universal da natureza*. Em primeiro lugar, o que Kant chama de *máxima* é o que nós estamos chamando de motivação, mais precisamente, um princípio subjetivo que explica uma ação. Se, por exemplo, eu decido pedir dinheiro emprestado sem a intenção de pagar ao meu credor, porque estou passando por um momento de necessidade, a máxima que explicaria minha ação seria algo do tipo: 'se for necessário, deve-se pedir dinheiro emprestado, mesmo sem a intenção de pagar a quem se deve'. Além disso, por "lei universal da natureza", Kant refere-se ao tipo de ordem natural, como a lei da gravidade, à qual estamos submetidos sem exceção. Tendo isso em mente, podemos entender o resto dessa formulação como apresentando o seguinte procedimento para avaliar a moralidade de qualquer ação: você deve, em primeiro lugar, identificar qual é a máxima da sua ação. Em segundo lugar, você deve

concebê-la como universal, de modo que todos deveriam proceder da mesma maneira. No nosso exemplo, isso envolve imaginar que, *se for necessário, todos deveriam pedir dinheiro emprestado, sem a intenção de pagar a quem se deve*. Em terceiro lugar, devemos transformar essa máxima universalizada em uma lei da natureza: *se for necessário, todos pedem dinheiro emprestado, sem a intenção de pagar a quem se deve*. Note que, nessa etapa, a expressão 'dever' em 'devem pedir' desaparece, restando apenas uma descrição, como se a máxima agora se tratasse de um fato: justamente, uma lei da natureza. Em quarto e último lugar, devemos conceber um mundo em que essa lei da natureza é válida desde sempre. Seria possível, então, ter a intenção de agir daquele modo (pedir dinheiro emprestado sem a intenção de pagar) e honestamente desejar habitar um mundo em que essa máxima universalizada vale desde sempre como uma lei da natureza? O diagnóstico kantiano é que isso seria *irracional*, porque, nesse mundo que imaginamos, não faz mais sentido manter promessas. Pedir dinheiro emprestado, mesmo com a falsa promessa de pagar de volta alguém a quem você deve, seria absurdo, porque todas as pessoas, desde sempre, estariam violando promessas desse tipo. Não seria sequer compreensível qual seria o conteúdo de uma promessa nesse mundo imaginado. É com base nesse raciocínio estritamente *a priori* que Kant aplica uma espécie de teste de coerência à máxima ou motivação de uma ação. Se a máxima falha nesse teste, então ela não é um dever moral.

O mesmo tipo de procedimento mostra, por exemplo, que é sempre moralmente condenável mentir — mesmo que a mentira seja um meio para as melhores consequências dos envolvidos. Imagine, por exemplo, que você escolhe mentir para sua avó porque não quer que ela se preocupe com você. Na deontologia kantiana, isso é errado porque, se todos mentissem o tempo todo, como se obrigados por uma lei da natureza, a própria ideia de um ato de fala, de comunicar ou expressar um pensamento, não faria sentido e, portanto, seria incoerente mentir e simultaneamente desejar um mundo em que a mentira para o bem dos envolvidos é uma lei universal da natureza.

A segunda formulação do imperativo categórico é a *Fórmula da Humanidade*, seja ela: *aja de tal modo que você sempre trate a humanidade, seja em você mesmo ou em outra pessoa, nunca como apenas um meio, mas sempre como um fim em si mesmo*. Essa formulação pode parecer equivalente à famosa "regra de ouro" do senso comum, segundo a qual você deve fazer aos outros o que espera que façam consigo. Mas, estritamente falando, essa não é uma equivalência, porque regra de ouro exigiria, por exemplo, de um masoquista que ele se comportasse como um sádico, imputando aos outros a dor que gostaria de sentir em si. A ideia kantiana na segunda formulação do imperativo categórico é muito mais refinada, e ela parte de duas suposições importantes: em primeiro lugar, que o conceito de humanidade implica racionalidade e, em segundo lugar, que a racionalidade é um fim em si mesmo — a humanidade deve, portanto, ser sempre respeitada como um fim em si mesmo, não apenas como um instrumento para a satisfação dos nossos prazeres. Nós, é claro, dependemos de pessoas como instrumentos para a nossa satisfação e para a manutenção do nosso bem-estar. Precisamos, por exemplo, de pessoas em uma longa cadeia de trabalho, desde as plantações e desde as distribuidoras de alimentos, até as pessoas que trabalham em mercados, para que a comida chegue à nossa mesa. Mas, além desse uso instrumental, devemos sempre respeitá-las como fins em si mesmo em virtude da sua humanidade. Ou seja, essa formulação expressamente rejeita que seja moralmente permissível usar outras pessoas *apenas* para fins externos, como o bem-estar dos envolvidos — é esse tipo de consideração que faz com que hesitemos em puxar a alavanca no Dilema do Bonde para matar uma pessoa e salvar outras cinco. Além disso, a segunda formulação do imperativo categórico refere-se não apenas ao tratamento que temos de outras pessoas, mas também de nós mesmos. Ela exclui, por exemplo, que um sujeito possa cometer suicídio porque é tremendamente infeliz, visto que ele estaria tratando a si mesmo como um mero instrumento para acabar com seu sofrimento.

Kant apresenta uma terceira formulação (embora não a expresse imperativamente como as demais), que recebe o nome de *Fórmula da Autonomia*, seja ela: *aja de modo que, por meio das suas máximas, você poderia ser um legislador de leis universais*. A ideia aqui, novamente, é que suas máximas poderiam tornar-se leis — mas, dessa nova perspectiva o foco passa a ser no legislador, não naqueles que estão sujeitos à lei (como é na primeira formulação).

É motivo de grande debate a questão sobre se essas formulações são de fato equivalentes (Kant afirma que sim, mas sem explicar em que exatamente consiste essa equivalência), o que envolve saber se as três formulações sancionam (e proíbem) o mesmo conjunto de ações. A proibição da mentira, por exemplo, parece seguir-se de todas as formulações, mas não é claro que eu tenho o *dever* de ajudar os outros de acordo com a segunda formulação, que apenas proíbe o desrespeito à humanidade, enquanto esse dever parece seguir-se mais claramente da primeira formulação (como o próprio Kant argumenta). De qualquer modo, considerações deontológicas como as da ética kantiana servem de fundamento para os *Direitos Humanos Universais*, porque explicitam que é moralmente condenável usar uma pessoa (torturar, perseguir, sacrificar) para o bem-estar de outras, como fica claro na segunda formulação. Trata-se de uma ética profundamente otimista, pois pressupõe que a razão seja uma faculdade confiável e homogeneamente compartilhada, o que permitiria que chegássemos a acordos morais por nossa própria conta, sem depender de legisladores externos, como mandamentos divinos. De fato, é uma teoria bastante sofisticada, mas isso obviamente não é garantia de que esteja correta. Há várias objeções bem conhecidas que mostram que a aplicação dos procedimentos propostos por Kant gera resultados que não se afinam com nossas intuições morais.

Note, em primeiro lugar que, se mentir é sempre, invariavelmente errado (de acordo com a aplicação da primeira formulação do imperativo categórico), então é errado mentir para proteger uma pessoa inocente. Imaginemos que você recebe na sua casa uma pessoa de uma minoria que é perseguida em uma ditadura sanguinária. Ima-

gine que, em seguida, a polícia de estado passa procurando por essa pessoa, e você tem toda a certeza de que ela será torturada e assassinada. Segundo Kant, você não pode mentir que não a viu, ou que não está protegendo essa pessoa na sua casa. Essa objeção foi originalmente apresentada por Benjamin Constant (no seu exemplo original, tratava-se de um assassino perseguindo um inocente, mas a catástrofe humanitária que foi o nazismo para sempre modificou o exemplo e lhe deu novos tons). Em resposta, Kant escreveu um curto ensaio em que, por assim dizer, "mata no peito" e mantém a consistência com a sua posição na *Fundamentação,* alegando que "ser honesto em todas as declarações é, portanto, um comando sagrado incondicional da razão, e que não pode ser limitado por quaisquer conveniências".

Pode-se argumentar também que a ética kantiana é ao mesmo tempo excludente e permissiva demais (por incrível que pareça). Kant, por exemplo, não usaria o Tinder — pois, afinal de contas, o sexo casual parece não passar de um uso meramente instrumental das pessoas envolvidas (ainda que mutuamente instrumental), e isso não nos parece moralmente condenável, não obstante os pudores kantianos. Desse modo, essa teoria moral parece ser excludente demais, pois ela proíbe o que não diríamos ser inerentemente condenável. Por outro lado, sua teoria parece permitir atos absurdos, pois a segunda formulação do imperativo categórico é baseada na racionalidade que é implicada pela humanidade. Note, no entanto, que animais não-humanos não seriam racionais no sentido que Kant (e maioria da tradição filosófica) entende o conceito de racionalidade, isto é, como uma capacidade deliberativa que envolve o uso da linguagem. Ou seja, animais não-humanos não têm a característica que nos marca como fins em nós mesmos, como merecedores de respeito irrevogável. Desse modo, a ética kantiana sancionaria o uso de animais como meros instrumentos para nossa diversão — torturando-os, por exemplo — porque não são seres dotados de racionalidade. Isso é claramente absurdo. Uma resposta kantiana poderia em princípio revisar o que se entende por racionalidade, de modo a adotar uma concepção mais inclusiva, uma que contemple, em diferentes graus,

animais não humanos. O problema é que isso implicaria revisar também toda teoria epistemológica kantiana. Portanto, na medida em que essa resposta seria mais onerosa do que o desejável, é claro que a ética kantiana, além de excludente demais, é permissiva demais.

Talvez a mais contundente objeção à ética kantiana seja baseada na idealização da racionalidade humana. Mesmo colocando de lado a possibilidade de que a racionalidade não seja uma faculdade com fundamentos independentes da experiência, de modo que pessoas poderiam exibir diferentes níveis e até mesmo diferentes tipos de racionalidade (o que Kant de qualquer modo não aceitaria), resta a dúvida de se é ao menos possível aplicar um procedimento tão complexo como o que nos manda o imperativo categórico em cada uma de suas formulações. O procedimento pode fazer sentido se pensarmos em dilemas morais dramáticos, mas não é óbvio que seja aplicável em questões morais de dia-a-dia que exigem respostas rápidas. Pelo contrário: parece mais intuitivo que uma pessoa que, como dizemos "para pra pensar" se deve agir de um modo ou de outro diante de um acontecimento qualquer já está agindo de um modo minimamente incorreto. Talvez uma boa decisão moral seja aquela que é tomada *sem pensar*, seja no conteúdo do dever ou nas consequências da ação, uma espécie de manifestação natural e irrefletida de bom-caráter e virtude.

Ética e política

Muito frequentemente as questões éticas são associadas às questões políticas. Isso não é por acaso, pois ambas dizem respeito a um aspecto inequivocamente normativo da conduta humana, isto é, ao que deve ser feito. Notemos, no entanto, que o vocabulário que empregamos para tratar de questões morais e de questões políticas é diferente: quando falamos de ética, usamos termos como 'correto', 'certo', 'bom' e 'dever' (e suas contrapartes negativas), mas, ao falarmos de política, a ideia central ao redor da qual orbitam todas as outras é a de *justiça*. Há, intuitivamente, uma diferença importante entre os

conceitos de *certo* e de *justo*, pois a questão sobre se uma ação é correta direciona o enfoque para uma avaliação individual, da ação *de um agente*. Diferentemente, a questão sobre se uma ação é justa remete a uma dimensão *pública*, que envolve instituições sociais, políticas e econômicas. A questão da distribuição e riquezas, por exemplo, é uma questão de justiça (ou injustiça), não exatamente, ou talvez não apenas, uma questão de certo (ou errado).

Ambas as teorias morais apresentadas acima, o consequencialismo e a deontologia, são fonte de inspiração para teorias políticas. O utilitarismo de Bentham já oferecia originalmente um critério para a aceitabilidade de decisões políticas, a saber, a ponderação do bem-estar dos afetados, de modo que uma decisão política é justa apenas na medida em que maximiza o bem-estar dos envolvidos. Uma lei, por exemplo, que aumenta o sofrimento dos envolvidos para o benefício de uns poucos seria claramente injusta. É claro, as mesmas objeções à ética utilitarista ocorrem quando tratamos de uma teoria da justiça utilitarista, pois seria possível, na versão clássica, que pessoas sejam submetidas a intenso sofrimento para maximizar o bem-estar de outras. Isso, no entanto, intuitivamente não configura uma situação política justa. Para Bentham, o bem-estar é em última análise uma função de prazeres e sofrimentos, mas utilitaristas contemporâneos por vezes substituem o hedonismo por interpretações menos reducionistas, como o aperfeiçoamento humano (a busca pela excelência) ou a preservação de direitos.

Alguns elementos da ética kantiana são fundamentais para a filosofia política desenvolvida pelo filósofo norte-americano John Rawls a partir dos anos 70. Como mencionamos no começo, a ética, como área de investigação filosófica, havia sido marginalizada da filosofia analítica durante boa parte do século XX, e o mesmo ocorreu com a filosofia política, que era vista como uma temática característica da tradição filosófica desenvolvida no continente europeu. O panorama mudou com a publicação com o livro de Rawls *Uma Teoria da Justiça*, em 1971, que reviveu o interesse pela filosofia política — que, a partir

de então, passou a ser tratada com as ferramentas típicas da filosofia analítica.

Naquele livro, Rawls retoma uma estratégia argumentativa típica da filosofia política moderna, a saber, o experimento mental do *contrato social*. O contratualismo moderno, especialmente aquele desenvolvido pelo filósofo britânico Thomas Hobbes no século XVII, surge como uma tentativa de explicar a origem da autoridade política. Na versão hobbesiana do experimento, as pessoas habitavam um estado de natureza sem quaisquer restrições anteriormente ao estabelecimento de convenções sociais. Hobbes supõe que, em tal estado de natureza, não existem valores éticos, de tal modo que o comportamento humano seria predatório, estando todos contra todos. Disso se segue que as pessoas no estado de natureza estariam racionalmente compelidas a abrir mão da sua liberdade irrestrita (uma liberdade negativa caracterizada pela ausência de impedimentos) para preservar sua vida. Essa concessão seria institucionalizada pelo contrato social, que conferiria todo o poder de decisão política à figura de um soberano. Mais tarde, John Locke, outro filósofo britânico, desenvolve outra influente versão do contrato social. Na sua versão, o contrato social é firmado para preservar, além do direito natural à vida (um direito que todas as pessoas possuem independentemente de fazerem parte de uma sociedade), também os direitos naturais à propriedade e à liberdade. O resultado é o estabelecimento de direitos positivos que preservam e ampliam os direitos naturais — ou seja, o contrato social tem como finalidade, entre outras, preservar a propriedade.

Os argumentos do contratualismo moderno são bem conhecidos, e aqui é importante tê-los em mente apenas para compará-los com o argumento de inspiração contratualista que Rawls apresenta. Notemos em especial que as circunstâncias que caracterizam a passagem do estado de natureza para o estado social, de acordo com os modernos, são tais que o uso da racionalidade é estritamente instrumental. No experimento moderno, as pessoas que teriam sua vida ameaçada, no caso de Hobbes, ou também sua propriedade e sua liberdade, no caso de Locke, raciocinariam que o melhor meio para preservar

aqueles bens seria abrir mão da sua liberdade negativa irrestrita. Seria, portanto, um raciocínio instrumental que dá origem ao contrato social no contratualismo moderno. É claro, no entanto, que uma organização política em que toda autoridade está concentrada em uma mesma figura não é justa. É menos claro, mas é argumentável, que uma organização política que preserva a propriedade privada independentemente do modo como ela foi adquirida não é justa. O experimento mental de Rawls evita esse tipo de resultado que desafia nossas intuições sobre justiça porque (supostamente) exibe em um uso *constitutivo* da racionalidade, de modo que a estrutura social resultante do contrato seria justa *porque* seria racional.

Especificamente, o experimento mental de Rawls não consiste na suposição de um estado de natureza em que pessoas seriam irrestritamente livres, mas em um cenário hipotético que ele chama de *posição original*. A posição original é uma situação em que as pessoas decidiriam conjuntamente quais seriam as características estruturais fundamentais de uma sociedade justa. É claro, a posição original não é um evento histórico, ela nunca ocorreu, e esse experimento não é utilizado para explicar as origens históricas de uma estrutura social vigente. O objetivo aqui é oferecer as razões pela escolha de uma estrutura social mínima justa (*mínima* porque, a partir dessa decisão pela estrutura básica da sociedade, outros aspectos podem ser modificados e reformados de acordo com contingências posteriores).

Uma sociedade justa, Rawls supõe, consistiria minimamente em uma distribuição justa de bens primários, isto é, bens que os indivíduos preferem ter mais do que ter menos, como direitos, liberdades, oportunidades, rendas e riqueza — bens que são finitos e, portanto, não podem ser distribuídos ao máximo para todos. É claro, esse contrato não pode ser negociado com base em fatores arbitrários, como desequilíbrios de poder, capacidades de persuasão e diferenças cognitivas dos participantes. Rawls então sugere que as pessoas na posição original sejam concebidas como que sob o *véu da ignorância*. Nessa circunstância hipotética, os contratantes seriam ignorantes acerca de questões contingentes da sua existência, como suas prefe-

rências, seus dons, suas habilidades e disposições, sua estrutura física, sua pertença a uma classe social, suas riquezas, seu gênero e sua etnia. Desse modo, não teriam nada além da sua capacidade racional pura (e de algumas informações adicionais e bastante gerais, como a finitude dos bens a serem distribuídos e o fato de que as pessoas têm diferentes planos de vida, por exemplo). Portanto, Rawls supõe — de modo tipicamente kantiano — que seja possível separar a racionalidade de contingências empíricas. Esse é um ponto que pode ser explorado para criticar a teoria rawlsiana a partir de uma perspectiva naturalizada, segundo a qual a racionalidade estaria em continuidade com nossas outras capacidades cognitivas. Com efeito, não parece haver um motivo para manter que a racionalidade possa ser completamente separada do mundo empírico, a não ser certo preconceito apriorístico. De qualquer modo, é a partir da deliberação de seres estritamente racionais (em um cenário imaginário) que seria acordada uma estrutura social mínima justa, caracterizando um uso constitutivo da razão pela decisão de princípios de justiça que qualquer um, em condições iguais, aceitaria.

Qual seria, então, o resultado de um contrato aceito sob o véu da ignorância? Ou seja, qual seria a estrutura social mínima justa que podemos pensar a partir da racionalidade apenas, não das nossas preferências pessoais? Em primeiro lugar, as pessoas na posição original não escolheriam uma organização social que permitisse a privação de alguns bens primários a um conjunto de pessoas (pessoas de certa classe, ou de certa etnia, ou de certa orientação sexual, por exemplo), porque os contratantes não teriam acesso à informação sobre se eles mesmos seriam prejudicados com essa distribuição. Portanto, os contratantes aceitariam como um princípio fundamental o que Rawls chama de *princípio da liberdade igual*: todas pessoas devem ter as mesmas liberdades e direitos básicos, e esse conjunto de liberdades e direitos básicos não podem ser violados para a obtenção de outros bens sociais. Pense do seguinte modo: se você tivesse que dividir um bolo entre 10 pessoas, você não cortaria um pedaço maior do que os outros 9 porque, numa situação de véu da ignorância, você

não saberia qual é o seu pedaço. Seria racional que você cortasse o bolo em pedaços iguais.

Em segundo lugar, os contratantes aceitariam o que Rawls chama de *princípio de diferença*, segundo o qual desigualdades sociais e econômicas são permissíveis apenas na medida em que representam uma vantagem para todos e favorecem os menos favorecidos da sociedade. Esse é um princípio muitas vezes apresentado como motivação para políticas de ações afirmativas. Para explorar um pouco mais a metáfora do bolo, imagine que uma das 10 pessoas está faminta enquanto os outros estão bem alimentados, de modo que seria aceitável que ela recebesse um pedaço um pouco maior do bolo. Agora imagine que, ao partilhar o bolo de modo que a pessoa faminta ganhe um pedaço maior, uma pessoa bem alimentada acabaria ganhando um pedaço um pouco maior também. Não parece haver nada de injusto nessa distribuição. Na concepção liberal de sociedade que Rawls propõe, portanto, uma distribuição de riquezas que permite a pessoas de uma classe acumular mais riqueza do que as pessoas de outras classes não seria uma distribuição injusta se fosse o caso que as pessoas menos favorecidas também ganhassem com o acúmulo de riqueza pelas outras classes. Com efeito, essa seria uma sociedade *mais justa* do que aquela em que uma distribuição estritamente igualitária de riquezas fizesse com que todas as pessoas fossem igualmente pobres.

Juntos, esses dois princípios constituem a base do que Rawls chama inicialmente de *justiça como equidade*. Essa é uma ideia é muito persuasiva, e ela foi muito influente além da filosofia política — Rawls, no entanto, foi também muito criticado em vários pontos da sua argumentação, como ao implicar que seria justo taxar riquezas acumuladas se isso favorecesse os menos favorecidos. A crítica geralmente procede constatando que isso violaria direitos individuais (um ponto levantando por Nozick). De qualquer modo, Rawls foi um dos principais responsáveis por uma revolução na filosofia analítica, não apenas por efetivamente criar uma tradição contratualista contemporânea, mas também por resgatar o prestígio perdido da filosofia política e torná-la uma área fértil e independente de arcaísmos mar-

xistas que dominavam a filosofia política europeia e que nós *ainda* temos dificuldades de superar em solo brasileiro.

Considerações finais

As perplexidades filosóficas em ética são talvez as mais próximas das nossas questões de dia-a-dia, e o modo como filósofos pretendem respondê-las, isto é, através da concepção de teorias morais, envolve o teste com o que nós diríamos sobre casos reais ou imaginários. A consulta às nossas intuições, no entanto, mostrou que não há consenso sobre como melhor entender os conceitos de *bom, certo, justo,* etc. Isso pode levar a uma espécie de *ceticismo moral,* segundo o qual simplesmente não haveria significado para aqueles conceitos. Mas essa não é a única reação possível a essa dificuldade, pois da discordância sobre qual é a melhor teoria disponível para explicar certo fenômeno (nesse caso, nossos usos de conceitos éticos) não se segue que todas as teorias são falsas. Outra possibilidade, talvez mais atraente, consiste em conceder que as melhores teorias éticas (as que melhor resistem ao teste das nossas intuições) na melhor das hipóteses apenas capturam diferentes dimensões ou facetas dos nossos conceitos morais, e que talvez teorias diferentes sejam mais ou menos adequadas para lidar com tipos específicos de situação — o utilitarismo seria preferível para lidar com o caso da mentira, um tipo de caso com que a deontologia kantiana é incapaz de lidar, enquanto uma ética dos deveres seria preferível para lidar com o caso do desrespeito à vida, um tipo de caso que apresenta uma objeção contundente ao utilitarismo.

Leituras recomendadas

Sandel (2015) apresenta uma claríssima introdução à ética e à filosofia política. Para excertos de obras originais comentados, veja os capítulos quinto e sexto de BonJour e Baker (2010), que incluem discussões sobre relativismo, consequencialismo, deontologia, contratualismo clássico e contemporâneo. O décimo primeiro ca-

pítulo de Rachels (2009) trata de modo muito claro a questão da objetividade da ética. A interpretação da primeira formulação do imperativo categórico é inspirada na reconstrução feita por Rawls (2006). Para mais sobre a teoria moral de Kant, veja Sober (2006) e, para a teoria da justiça de Rawls, veja Vaz (2006).

Coda

Agora que chegamos ao fim do nosso percurso, podemos voltar à questão metafilosófica com que começamos: *o que é a filosofia?* Nós apresentamos duas respostas a essa pergunta, a saber, a concepção tradicional de filosofia como disciplina de segunda ordem e a concepção naturalizada de filosofia como uma investigação em continuidade com as ciências empíricas. Ao explorarmos essas duas visões, nós efetivamente *fizemos filosofia* — apresentamos e examinamos criticamente respostas a algumas de nossas perplexidades mais fundamentais. Apesar de que isso não nos permita resolver definitivamente nossa questão inicial, o modo como conduzimos nossa atividade nos diz algo a respeito do *significado da prática filosófica*. Pois, embora cínicos denunciem a filosofia como uma tarefa inútil, como uma perda de tempo que não teria lugar neste mundo, o que nossas discussões mostram é que o exercício filosófico é um ato de resistência em defesa da racionalidade. Essa é a verdadeira razão pela qual a filosofia é tão frequentemente vilificada: fazer filosofia é um entrave ao projeto de esvaziamento da razão humana. Fazer filosofia não é uma perda de tempo, pelo contrário, é uma ameaça àquele projeto. Por isso continuamos.

Referências bibliográficas

BonJour, L., & Baker, A. (2010). *Filosofia: textos fundamentais comentados* (2a ed.). Porto Alegre: Artmed.

Branquinho, J., Murcho, D., & Gonçalves Gomes, N. (Eds.). (2006). *Enciclopédia de termos lógico-filosóficos*. São Paulo: Martins Fontes.

Bruce, M., & Barbone, S. (2013). O problema do mal. In M. Bruce & S. Barbone (Eds.), *Os 100 argumentos mais importantes da filosofia ocidental*. São Paulo: Cultrix.

Castro, E. (2014). Realismo/Anti-Realismo. In J. Branquinho & R. Santos (Eds.), *Compêndio em Linha de Problemas da Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa.

Castro, E., & Fernandes, D. (2014). O problema da indução. In J. Branquinho & R. Santos (Eds.), *Compêndio em Linha de Problemas da Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa.

Casullo, A. (2015). Conhecimento a priori. *Investigação Filosófica, E4*(2), 86–133.

Chakravartty, A. (2017). Realismo e antirrealismo; metafísica e empirismo. *Investigação Filosófica, 8*(1).

Chalmers, A. F. (1993). *O que é ciência afinal?* São Paulo: Brasiliense.

Costa, C. F. (2005). Teorias da Verdade. *Crítica na Rede*. https://criticanarede.com/met_tverdade.html

Coura, M. (2014). Descrições Definidas. In *Compêndio em Linha de Problemas da Filosofia Analítica*. Centro de Filosofia da Universidade de Lisboa.

Dinucci, A., & Duarte, V. (2016). *Introdução à Lógica Proposicional Estoica*. São Cristóvão: UFS.

Ferreira, S. S. (2010). Introdução à teoria das descrições de Russel. *Crítica na Rede*. https://criticanarede.com/descricoes.html

Glock, H.-J. (2011). *O que é a filosofia analítica?* Porto Alegre: Artmed.

Hansson, S. O. (2018). Ciência e pseudo-ciência. *Crítica na Rede*.

Hegenberg, L., & Silva, M. F. de A. e. (2005). *Novo dicionário de lógica*. Rio de Janeiro: Pós-Moderno.

Hetherington, S. (2013). O argumento cartesiano do sonho sobre o mundo exterior. In M. Bruce & S. Barbone (Eds.), *Os 100 argumentos mais importantes da filosofia ocidental*. São Paulo: Cultrix.

Jordan, W. (2016). O Nascimento da Filosofia. *Crítica na Rede*. https://criticanarede.com/hist_origemfilo.html

Kahneman, D. (2011). *Rápido e devagar: duas formas de pensar*. Rio de Janeiro: Objetiva.

Kant, I. (2018). *Crítica da Razão Pura*. (F. C. Mattos, Trand.) (4a ed.). Petrópolis: Vozes.

Kneale, W., & Kneale, M. (1991). *O desenvolvimento da lógica*. Lisboa: Fundação Calouste Gulbenkian.

Kripke, S. (2012). *O nomear e a necessidade*. (R. Santos & T. Filipe, Trans.). Lisboa: Gradiva.

Lazier, J., & Gaul, B. (2013). O argumento do Cogito de Descartes e Agostinho. In M. Bruce & S. Barbone (Eds.), *Os 100 argumentos mais importantes da filosofia ocidental*. São Paulo: Cultrix.

Mortari, C. (2001). *Introdução à lógica*. São Paulo: UNESP.

Motloch, M. (2016). Necessidade. In J. Branquinho & R. Santos (Eds.), *Compêndio em Linha de Problemas da Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa.

Moulines, C. U. (2020). *O desenvolvimento moderno da filosofia da ciência (1890–2000)*. (C. Abreu, Trad.). São Paulo: Associação Filosófica Scientiae Studia.

Murcho, D. (2015). Uma introdução ao necessário a posteriori. *Crítica na Rede*. https://criticanarede.com/introkripke.html

Murcho, D. (2019). É lógico, Aristóteles. *Crítica na Rede*.

Nolt, J., & Rohatyn, D. (1991). *Lógica*. São Paulo: McGraw-Hill.

O'Brien, D. (2013). *Introdução à teoria do conhecimento*. Lisboa: Gradiva.

Papineau, D. (2011). A epistemologia da ciência. *Crítica na Rede*. https://criticanarede.com/episciencia.html

Rachels, J. (2009). *Problemas da Filosofia*. Lisboa: Gradiva.

Rawls, J. (2006). A primeira formulação do imperativo categórico de Kant. *Crítica na Rede*. https://criticanarede.com/eti_kantrawls.html

Rolla, G. (2018). *Epistemologia: uma introdução elementar*. Porto Alegre: Editora Fi.

Ruffino, M. (2013). Contingente a priori. In J. Branquinho & R. Santos (Eds.), *Compêndio em Linha de Problemas da Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa.

Russell, B. (2008). *Os Problemas da Filosofia*. Coimbra: Edições 70.

Sandel, M. (2015). *Justiça*. (H. Matias & M. A. Máximo, Trans.). Rio de Janeiro: Civilização Brasileira.

Silveira, F. L. da. (2002). *Caderno Brasileiro de Ensino de Física, 19*(Especial), 28–51.

Sinclair, R. (2013). Os dois dogmas do empirismo de Quine. In M. Bruce & S. Barbone (Eds.), *Os 100 argumentos mais importantes da filosofia ocidental*. São Paulo: Cultrix.

Sober, E. (2006). A teoria moral de Kant. *Crítica na Rede*. https://criticanarede.com/eti_kant.html

Teixeira, C. (2014). Conhecimento a priori. In J. Branquinho & R. Santos (Eds.), *Compêndio em Linha de Problemas da Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa.

Teixeira, C. (2015). Analiticidade. In J. Branquinho & R. Santos (Eds.), *Compêndio em Linha de Problemas da Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa.

Vaz, F. (2006). A teoria da justiça de John Rawls. *Crítica na Rede*. https://criticanarede.com/pol_justica.html

Williamson, T. (2019). *Filosofar: da curiosidade comum ao raciocínio lógico*. Lisboa: Gradiva.

Williges, F. (2007). Idealismo Transcendental e Ceticismo em Kant. *Barbarói*, (26), 117–132.